



Contributions à l'analyse asymptotique des systèmes multi-antennes

Cédric Artigue

► To cite this version:

Cédric Artigue. Contributions à l'analyse asymptotique des systèmes multi-antennes. Traitement du signal et de l'image. Université Paris-Est, 2010. Français. <tel-00865714>

HAL Id: tel-00865714

<https://tel.archives-ouvertes.fr/tel-00865714>

Submitted on 25 Sep 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Thèse
présentée
devant l'Université de PARIS EST
pour obtenir
le grade de DOCTEUR DE L'UNIVERSITÉ DE PARIS
EST
Mention TRAITEMENT DU SIGNAL ET DES IMAGES
par

Cédric ARTIGUE

Équipe d'accueil : Signal et Communications
École Doctorale : MSTIC
Composante universitaire : LABORATOIRE D'INFORMATIQUE DE L'INSTITUT
GASPARD MONGE

Titre de la thèse :
**Contributions à l'analyse asymptotique
des systèmes multi antennes**

Soutenue le 2 juin 2010 devant la commission d'examen.
Composition du jury :

Président

Dirk	SLOCK	Professeur à Eurecom
------	-------	----------------------

Rapporteurs

Walid	HACHEM	Chercheur CNRS
David	GESBERT	Professeur à Eurecom
Merouane	DEBBAH	Professeur à Supelec

Examineurs

David	DECLERCQ	Professeur à l'ENSEA
Philippe	FORSTER	Professeur à l'ENS Cachan
Philippe	LOUBATON	Professeur à l'Université de Paris Est (Directeur de thèse)

Remerciements

Je remercie tout d'abord ma famille ainsi que mes amis pour leur soutien inconditionnel lors de ma thèse.

Je tiens également à remercier mon directeur de thèse autant pour sa grande rigueur que pour sa sympathie.

Je remercie finalement l'ensemble des membres de mon jury.

Table des matières

1	Introduction	11
1.1	Contexte de la thèse	11
1.2	Plan	15
1.3	Contributions	17
I	MIMO-CDMA	19
2	Modélisation des signaux	21
2.1	Description des signaux émis	23
2.1.1	Étalement des données	25
2.1.2	Allocation de puissance	26
2.1.3	Embrouillage (Scrambling)	27
2.1.4	Mise en forme	28
2.2	Modélisation du canal de propagation	28
2.2.1	Modèle à temps continu	30
2.2.2	Modèle à temps discret	30
2.2.3	Modèle aléatoire gaussien	32
2.3	Modélisation réaliste du réseau	33
2.3.1	Géométrie du réseau	33
2.3.2	Niveaux de puissance	36
2.4	Signal reçu	37
2.4.1	Expression du signal reçu	37
2.4.2	Rapport signal à bruit réel	38
3	Démodulation	39
3.1	Le filtre de Wiener optimal	39

3.2	Le filtre de Wiener sous-optimal	45
4	Analyse de performance	47
4.1	Indice de performance	47
4.2	Calcul du SINR	48
4.3	Première approximation	54
4.4	Seconde approximation	57
4.5	Simulations	59
4.5.1	Présentation des scénarios de simulation	60
	Scénario 1 : Aucune interférences entre cellules	60
	Scénario 2 : A l'intérieur de la cellule	61
	Scénario 3 : A l'intersection de 2 cellules	63
	Scénario 4 : A l'intersection de 3 cellules	65
4.5.2	Validation des approximations	67
	Allocation de puissance uniforme	67
	Allocation de puissance non uniforme	71
5	Contrôle de puissance à long terme	77
5.1	Idée générale	77
5.2	Modèle des signaux	79
5.3	Démodulation	80
5.3.1	Performances du filtre de Wiener sous-optimal	80
	Première approximation du SINR	81
	Seconde approximation du SINR	81
5.3.2	Détermination de l'allocation de puissance à long terme	82
	Faisabilité de la méthode dans un contexte réel	83
5.4	Efficacité de la méthode d'allocation de puissance à long terme	84
5.4.1	Contrôle de puissance associé au WCDMA	84
5.4.2	Résultats de simulation	85
II	Optimisation de précodeurs	89
6	Avant-Propos	91

7	Article présenté à IEEE Globecom 2008	97
7.1	Abstract	99
7.2	Introduction	99
7.3	Asymptotic expression of \mathcal{C}_{mmse}	100
7.4	Design of optimal precoders.	103
7.4.1	Influence of the transmit correlation matrix \mathbf{C} on $\bar{\mathcal{C}}_{mmse}$	103
7.4.2	Structure of linear precoders.	104
7.4.3	Study of Problem 1.	106
7.4.4	Numerical illustration of the precoder optimization.	107
7.5	Concluding remarks.	108
8	Article soumis à IEEE Information Theory	111
8.1	Introduction	113
8.2	Problem statement.	117
8.3	Derivation of the large system approximation of I_{mmse}	118
8.4	Structure of optimal precoders.	125
8.5	Maximization of \bar{I}_{mmse}	131
8.5.1	Maximization of \hat{I}_{mmse} in the case of i.i.d. channels.	131
8.5.2	Study of Problem 5.	136
8.5.3	Numerical illustration	136
8.6	Concluding remarks.	138
8.7	Proof of Proposition 3.	138
8.8	Proof of Lemma 1.	149
9	Conclusion	151

Table des figures

2.1	Zones d'interférences	22
2.2	Chaine d'émission	24
2.3	Procédure d'étalement	25
2.4	Procédure d'allocation de puissance	26
2.5	Procédure d'embrouillage	27
2.6	Environnement de propagation à trajets multiples	29
2.7	Réseau constitué de cellules hexagonales	35
4.1	Scénario 1, géométrie du réseau	60
4.2	Scénario 2, géométrie du réseau	62
4.3	Scénario 3, géométrie du réseau	64
4.4	Scénario 4, géométrie du réseau	66
4.5	Scénario 1, puissance uniforme	69
4.6	Scénario 2, puissance uniforme	69
4.7	Scénario 3, puissance uniforme	70
4.8	Scénario 4, puissance uniforme	70
4.9	Scénario 1, Distribution des puissances non uniforme	74
4.10	Scénario 2, Distribution des puissances non uniforme	74
4.11	Scénario 3, Distribution des puissances non uniforme	75
4.12	Scénario 4, Distribution des puissances non uniforme	75
5.1	BER moyen versus distance mobile-Station de base	86
7.1	Average of the relative error.	103
7.2	Capacity with and without optimum precoding.	107
8.1	Accuracy of the large system approximant	126
8.2	Relative error	126

8.3	Relevance of the antenna selection scheme, $s = 6$ versus $s = 8$	135
8.4	Relevance of the antenna selection scheme, SNR = 15 dB	136
8.5	Impact of precoding scheme	137

Liste des tableaux

2.1	Paramètres de propagation	36
4.1	Scénario 2 : Distances Mobile - Stations de Base	63
4.2	Scénario 3 : Distances Mobile - Stations de Base	63
4.3	Scénario 4 : Distances Mobile - Stations de Base	65

Chapitre 1

Introduction

1.1 Contexte de la thèse

Cette thèse s'est déroulée dans le cadre d'un contrat CIFRE entre l'université de Paris Est Marne la Vallée et Freescale Semiconductor sur son site basé à Toulouse. Le sujet principal a été l'étude de la liaison descendante de systèmes de télécommunication disposant de plusieurs antennes d'émission et de réception et équipés d'un récepteur de Wiener plus généralement appelé récepteur MMSE (Minimum Mean Squared Error).

Il est de nos jours bien connu que l'utilisation de plusieurs antennes en émission et réception apporte un gain substantiel à la capacité de Shannon associée aux systèmes de communications sans fil et améliore notablement leur efficacité spectrale. En règle générale, le canal de propagation MIMO situé entre le réseau de t antennes d'émission et le réseau de r antennes de réception est modélisé par une matrice \mathbf{H} de dimension $r \times t$ que nous appellerons matrice de canal MIMO. Dans un contexte d'utilisation réelle, la mobilité et la présence d'un grand nombre de trajets réfléchis entraînent une variation plus ou moins rapide de l'état du canal de propagation au cours du temps. Afin de faciliter l'étude de l'impact de ces phénomènes sur les performances du système considéré il est devenu tout à fait naturel d'assimiler l'évolution temporelle de la matrice de canal aux réalisations d'une matrice aléatoire. En considérant ce modèle aléatoire, Telatar ([1]) réalisa vers le milieu des années 90 que la capacité ergodique de Shannon associée au canal MIMO augmentait linéairement selon $\min(r, t)$ lorsque la puissance totale allouée aux différents flux transmis par les antennes d'émission était fixée. Plus précisément, en considérant que les entrées de

la matrice de canal \mathbf{H} sont des variables aléatoires centrées, indépendantes et identiquement distribuées (i.i.d.) et en notant $I(\rho) = \log \det \left(\frac{\rho}{t} \mathbf{H} \mathbf{H}^H + \mathbf{I}_t \right)$ la capacité du canal MIMO \mathbf{H} pour un rapport signal à bruit égal à ρ , il détermina que la capacité par antenne d'émission $\frac{I(\rho)}{t}$ converge vers une limite déterministe. Il utilisa pour cela un résultat classique issu de la théorie des matrices aléatoires ([2]) stipulant que la distribution des valeurs propres de la matrice de Gram $\mathbf{H} \mathbf{H}^H$ converge en probabilité vers une distribution déterministe lorsque t et r tendent vers l'infini en conservant un rapport $\frac{r}{t}$ constant strictement positif. Ces résultats permirent alors de valider la conjecture selon laquelle la capacité d'un système de communication MIMO augmentait linéairement avec le nombre d'antennes. De plus, la convergence se trouvant suffisamment rapide, ces résultats permirent de s'assurer de la pertinence de l'utilisation de la théorie des matrices aléatoires pour prédire les performances des systèmes de communication MIMO muni d'un nombre raisonnable d'antennes.

Naturellement, l'étape suivante fût alors d'appliquer les résultats de la théorie des matrices aléatoires afin d'étendre les résultats précédents à des modèles de canaux plus élaborés faisant intervenir un modèle de corrélation entre les différentes entrées de la matrice \mathbf{H} , traduisant un éventuel phénomène de corrélation spatial au niveau des réseaux d'antennes d'émission et de réception. L'un des modèles de corrélation les plus populaire est le modèle de Kronecker, pour lequel la matrice de canal MIMO est donnée par $\mathbf{H} = \mathbf{C}_R^{\frac{1}{2}} \mathbf{H}_{iid} \mathbf{C}_T^{\frac{1}{2}}$, où \mathbf{H}_{iid} est une matrice de dimension $r \times t$ dont les entrées sont centrées et i.i.d. et où les matrices \mathbf{C}_T de dimension $t \times t$ et \mathbf{C}_R de dimension $r \times r$ traduisent respectivement les phénomènes de corrélation spatiale au niveau des réseaux d'antennes d'émission et de réception. A l'aide de ce modèle, les auteurs de [3] montrèrent que la capacité ergodique par antenne d'émission $\frac{I(\rho)}{t}$ converge vers une limite déterministe solution d'une équation au point fixe. Plus tard, les auteurs de [4] déterminèrent la limite de $\frac{I(\rho)}{t}$ pour un modèle de corrélation plus général que celui de Kronecker.

La capacité ergodique de Shannon est une indice de performance de première importance lorsque le système de communication MIMO considéré est muni d'un décodeur au maximum de vraisemblance. Cependant, l'implémentation réelle de ce type de récepteur nécessitant une complexité de calcul trop importante, il est raisonnable de se limiter à l'utilisation d'un récepteur linéaire de type MMSE aux performances moindres mais dont l'implémentation est par contre beaucoup moins coûteuse. Dans cette optique, il est alors nécessaire d'en étudier la capacité ergodique

correspondante I_{MMSE} à SNR fixé, définie comme la somme sur les t antennes d'émission des termes $\mathbb{E}(\log(1 + \beta_j))$ où β_j représente le rapport signal à bruit plus interférences (SINR) en sortie du récepteur MMSE adapté au flux transmis par la j^{me} antenne ([5]).

En ce qui concerne les systèmes MIMO basés sur le CDMA, l'étude de leur efficacité spectrale en uplink a déjà été menée lorsque le nombre d'utilisateurs et le facteur d'étalement tendent vers l'infini en conservant un rapport constant en considérant des codes d'étalements i.i.d. ([6]). Le downlink des systèmes CDMA mettant à profit des codes orthogonaux et non pas i.i.d., il est intéressant d'en étudier la capacité ergodique lorsque le nombre d'utilisateurs et le facteur d'étalement tendent vers l'infini et lorsque les nombres d'antennes en émission et en réception tendent vers l'infini. Dans [7] nous avons ainsi présenté un approximant \hat{I}_{MMSE} en grande dimension (utilisateurs, facteur d'étalement et antennes d'émission et de réception) de la capacité ergodique de systèmes MIMO-CDMA dans un environnement de propagation sélectif en fréquence en considérant un modèle de canal de kronecker traduisant un éventuel phénomène de corrélation spatiale au niveau des antennes d'émission, des codes d'étalement orthogonaux et des séquences de scrambling i.i.d. différentes pour chacune des antennes d'émission. Nous y montrons que la capacité ergodique augmente avec le facteur de charge du système et prouvons que lorsque le canal de propagation n'est pas sujet aux phénomènes de corrélation, le meilleur schéma de transmission n'est pas forcément l'utilisation de l'ensemble des t antennes d'émission comme cela peut être le cas pour la capacité de Shannon, mais qu'il est préférable de recourir à un schéma de sélection d'antennes. Dans la première partie de ce manuscrit, nous généralisons ces résultats à l'ensemble d'un réseau MIMO-CDMA multi utilisateurs et multi stations de base en intégrant un modèle d'allocation de puissance réaliste inclus dans la matrice de covariance de l'émetteur. A l'aide de notre approximation \hat{I}_{MMSE} de la capacité ergodique I_{MMSE} , qui se trouve être parfaitement fiable pour un nombre raisonnable d'antennes et d'utilisateurs, nous proposons un nouveau schéma d'allocation de puissance visant à garantir à un ensemble d'utilisateur un taux moyen d'erreurs binaires cible.

En ce qui concerne les systèmes MIMO génériques, l'étude et la conception de précodeurs maximisant la capacité ergodique I_{MMSE} est une question de première importance car c'est un moyen de tendre vers les débits limites pouvant être transmis de façon sûr aux utilisateurs équipés d'un récepteur MMSE. Ce problème de maxi-

misation a ainsi pu faire l'objet de nombreux travaux, notamment lorsque l'état du canal de propagation (ie. la matrice \mathbf{H}) est parfaitement connue par l'émetteur et le récepteur (voir [8]). Cependant, cette hypothèse n'étant pas réaliste, nous allons considérer dans nos travaux que le canal de propagation, choisi selon le modèle de kronecker tel qu'il est introduit dans [9], est parfaitement connu du récepteur mais que seules les statistiques du second ordre le sont au niveau de l'émetteur. Un problème soulevé par la conception de précodeurs est la difficulté de maximiser l'expression I_{MMSE} qui se trouve être assez compliquée. Résoudre le problème de façon directe entraine la nécessité d'évaluer numériquement les valeurs et vecteurs propres des matrices de précodage optimal, ce qui peut se révéler très couteux en terme de complexité de calcul car cela nécessite l'emploi d'algorithmes itératifs employant d'intensives technique de simulations de Monte Carlo visant à calculer le gradient de la fonction de coût. De plus la convergence de ces algorithmes ne peut être garantie car I_{MMSE} n'est en général pas concave. Comme il a été déjà introduit précédemment en ce qui concerne la capacité de Shannon des systèmes MIMO ([10],[11],[12]), nous proposons de remplacer le problème de la maximisation de I_{MMSE} par celui de la maximisation de son approximation déterministe \hat{I}_{MMSE} , obtenue en régime asymptotique lorsque les nombres d'antennes t en émission et r en réception tendent vers l'infini en conservant un rapport constant.

Dans [13] nous avons étudié la capacité ergodique associée aux systèmes MIMO équipé du récepteur MMSE évoluant au sein d'un environnement sélectif en fréquence lorsque le canal de propagation est connu du récepteur mais que seule les statistiques du second ordre sont connues au niveau de l'émetteur. Nous en proposons un approximant asymptotique \hat{I}_{MMSE} que nous utilisons afin de déterminer la structure d'un précodeur optimal maximisant \hat{I}_{MMSE} dans le cas où les antennes de transmission sont sujettes à un phénomène de corrélation spatiale et nous y montrons que les vecteurs propres gauches de la matrice de précodage optimale correspondent à ceux de la moyenne des matrices de covariance associées aux différents trajets et que ses valeurs propres sont les solutions d'un certain problème d'optimisation. Les résultats de simulations que nous y présentons nous assurent que l'approximation \hat{I}_{MMSE} que nous obtenons est suffisamment fiable pour un faible nombre d'antennes et pour un nombre de trajets $L > 1$. Cependant nous avons pu remarquer que l'erreur d'approximation pouvait se révéler assez importante lorsque l'on se plaçait dans le contexte d'un canal de propagation plat. Ceci nous poussa à proposer un

nouvel approximant de la capacité ergodique plus fiable. Ainsi dans [14], dont nous rappelons les résultats en seconde partie de ce document, nous avons considéré un canal de propagation plat sujets aux phénomènes de corrélation à l'émission et la réception toujours selon le modèle de Kronecker. Dans ces travaux, nous établissons que le terme d'erreur relatif associé à \hat{I}_{MMSE} est un terme de l'ordre de $O\left(\frac{1}{t}\right)$, ce qui correspond à une vitesse de convergence relativement faible lorsque on la compare à celle de l'approximant en grande dimension de la capacité de Shannon, dont le terme d'erreur associé se trouve être un terme en $O\left(\frac{1}{t^2}\right)$ ([15],[16]). Nous proposons alors d'utiliser une approximation en grande dimension améliorée, notée \bar{I}_{MMSE} , qui fût tout d'abord introduite dans [17] pour le cas des canaux MIMO i.i.d. avant d'être généralisée indépendamment dans [18] à l'aide de la méthode des répliques dont la rigueur mathématique reste encore à prouver et dans [19] à l'aide de résultats issus de la théorie des matrices aléatoires. Nous montrons que l'erreur relative associée à \bar{I}_{MMSE} est un terme en $O\left(\frac{1}{t^2}\right)$, améliorant ainsi les prédictions de [19] ($O\left(\frac{1}{t^{\frac{3}{2}}}\right)$) et [18] ($o\left(\frac{1}{t}\right)$). Les résultats de simulations nous permettant de nous assurer de la pertinence \bar{I}_{MMSE} dans le cas des canaux de propagation plat, nous utilisons finalement \bar{I}_{MMSE} plutôt que I_{MMSE} dans le but de déterminer une matrice de précodage maximisant la capacité ergodique. Nous déterminons ainsi une structure de matrice de précodage dont les vecteurs propres gauches correspondent à ceux de la matrice de covariance à l'émetteur et que ses valeurs propres sont les solutions d'un problème d'optimisation.

1.2 Plan

La première partie de ce document s'intéresse à l'étude des performances du downlink des systèmes MIMO-CDMA équipés d'un récepteur MMSE. Cette partie se divise en plusieurs sous-parties de la façon suivante :

- Dans la première sous-partie nous décrivons la structure des signaux que nous considérons dans notre étude et qui s'inspire du mode FDD du WCDMA issu de l'UMTS. Plus précisément, nous y abordons les concepts d'étalement, d'embrouillage et d'allocation de puissance. Nous présentons ensuite le modèle de canal de propagation MIMO adopté dans notre étude. Finalement nous exposons la structure de réseau multi-cellulaire que nous considérons et aboutissons ainsi à l'expression du signal reçu au niveau des mobiles.

- La seconde sous-partie est dédiée aux démodulateurs. Nous y présentons le récepteur de Wiener sous-optimal permettant d'estimer les symboles émis par les différentes stations de bases. Constatant que celui-ci n'est pas réalisable en pratique nous introduisons alors une structure sous-optimale implémentable, le récepteur de Wiener sous-optimal.
- La troisième partie s'intéresse au calcul des performances du récepteur de Wiener sous-optimal. Nous y présentons l'expression exacte du SINR correspondant conformément au modèle multi-cellulaire précédemment exposé. Constatant alors la complexité de cette expression nous menons successivement deux études asymptotiques. En considérant tout d'abord que le nombre d'utilisateurs et le facteur d'étalement tendent vers l'infini dans un rapport constant, la première étude asymptotique nous mène à une première approximation du SINR indépendante des réalisations particulières des codes de scrambling. En considérant ensuite le cas pour lequel les nombres d'antennes en émission et en réception tendent vers l'infini en conservant un rapport constant, la seconde étude asymptotique nous mène finalement à la seconde approximation du SINR qui se trouve être complètement déterministe. Cette sous-partie s'achève par la présentation de résultats de simulations qui, à travers quatre scénarios qui se veulent réalistes, permettent de s'assurer de la fiabilité des approximations obtenues.
- La quatrième sous-partie constitue un exemple concret d'utilisation des expressions asymptotiques obtenues dans les sections précédentes. Nous y présentons ainsi un nouvel algorithme d'allocation de puissance à long terme visant à garantir à un groupe d'utilisateurs un taux d'erreur binaire moyen donné. Des résultats de simulations sont alors présentés afin de valider la méthode de contrôle de puissance développée dans cette sous-partie.

La seconde partie de ce manuscrit présente les travaux liés à l'étude des performances des systèmes MIMO génériques dans le cas des canaux corrélés. Elle s'organise de la façon suivante :

- Dans un premier temps, nous exposons les modèles des signaux et le modèle de canal de propagation considérés et rappelons les principaux résultats obtenus dans [13] et [14] en ce qui concerne les approximations en grande dimension \hat{I}_{MMSE} et \bar{I}_{MMSE} de la capacité ergodique réelle I_{MMSE} , notamment leur vitesse de convergence. Nous rappelons alors les structures des précodeurs

obtenus en maximisant \hat{I}_{MMSE} et \bar{I}_{MMSE} plutôt que I_{MMSE} .

- Le chapitre suivant est constitué par l'article de conférence [13]
- Le dernier chapitre est constitué par [14], un article soumis à la revue IEEE Transaction on Information Theory

1.3 Contributions

Les différents travaux menés au cours de ces trois années de thèse ont donné naissance aux publications suivantes :

- [7] C. Artigue, P. Loubaton, B. Mouhouche, On the performance of spatial-multiplexing MIMO-CDMA downlink receivers based on chip rate equalization and despreading, in IEEE International Conference on Acoustics, Speech and Signal Processing, Las Vegas, March 2008
- [20] B. Mouhouche, C. Artigue, Coding Strategies for CDMA Packet Data Networks with Reduced Rank Multiuser detection, ARQ and Packet Combining, in IEEE International Workshop on Signal Processing Advances in Wireless Communications, Recife, Brazil, July 2008
- [13] C. Artigue, P. Loubaton, On the Ergodic Capacity of Frequency Selective MIMO Systems equipped with MMSE Receivers : An Asymptotic Approach, In IEEE Global Communications Conference, New Orleans, USA, November 2008
- [19] C. Artigue, P. Loubaton, On the Ergodic Capacity and Precoder Design of Flat Fading MIMO Systems Equipped with MMSE Receivers, In IEEE International Symposium on Information Theory, Seoul, Korea, June 2009
- [14] C. Artigue, P. Loubaton, On the precoder design of flat fading MIMO systems equipped with MMSE receivers : a large system approach, In IEEE Transactions on Information Theory, December 2009

Première partie

MIMO-CDMA

Chapitre 2

Modélisation des signaux

Nous considérons un réseau cellulaire comportant Q stations de base disposant chacune de t antennes d'émission ainsi que d'un certain nombre de mobiles équipés de r antennes en réception. Nous envisageons alors une stratégie de multiplexage spatial de sorte que chacune des stations de base est chargée de transmettre t flux indépendants sur son réseau d'antennes d'émission, chacun de ces flux contenant K suites de symboles destinées à K utilisateurs différents. La propagation du signal émis par chacune des stations de base vers les différents mobiles s'effectue ainsi à travers un canal MIMO comportant t entrées et r sorties. Nous définissons alors par le terme "cellule" la zone de couverture de chacune des stations de base. La disposition du réseau que nous considérons, illustrée par la figure 2.1, permet d'envisager des zones d'interférences entre cellules dont l'impact sera pris en compte lors de l'évaluation des performances du récepteur linéaire équipant le mobile d'intérêt.

La méthode d'accès multiple mise en jeu dans notre étude étant le CDMA, nous allons présenter la chaîne de transmission liée à cette technique et ensuite détailler les opérations d'étalement et d'embrouillage, ce qui nous permettra par la même occasion d'introduire les notations de base propres à notre modèle. Nous présenterons ensuite le modèle de canal de propagation sélectif en fréquence tel que nous l'envisageons. Finalement nous définirons le signal reçu au niveau du réseau d'antennes du mobile d'intérêt. Dans un souci de réalisme, le choix des différents paramètres sera effectué dans l'optique de rapprocher notre modèle de ce que l'on peut retrouver dans un réseau cellulaire réel basé sur le Frequency Division Duplex Wideband CDMA (FDD WCDMA), la technologie 3G européenne actuellement déployée, telle qu'elle est définie notamment dans [21] et [22] par l'organisme normalisateur 3GPP.

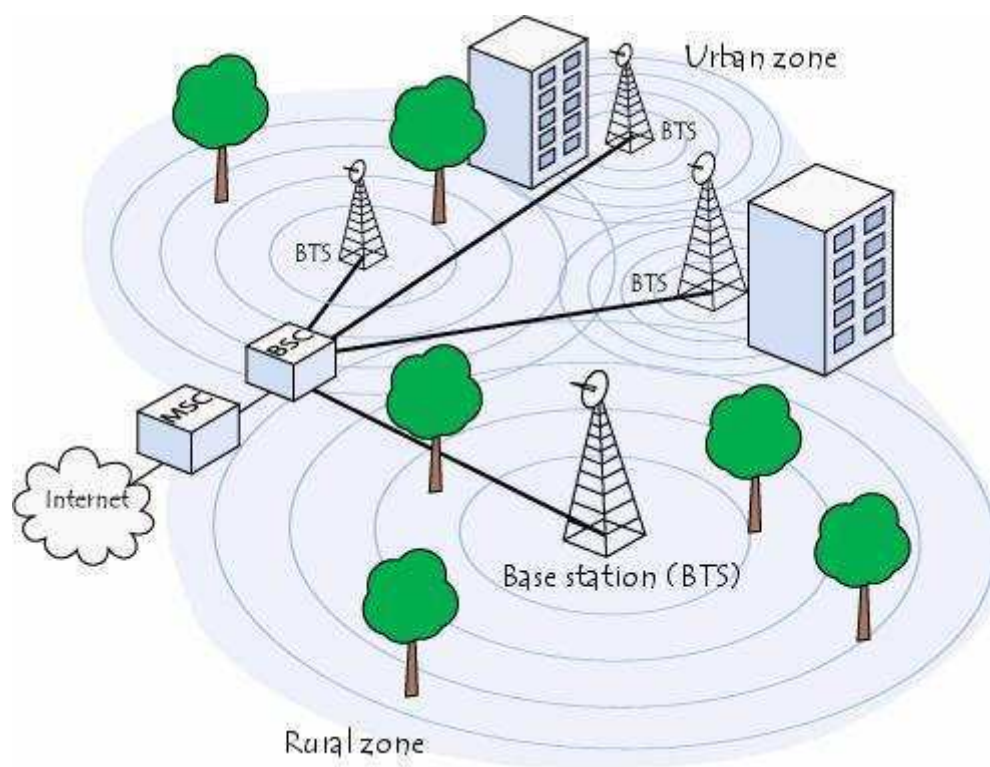


FIGURE 2.1 – Zones d'interférences

2.1 Description des signaux émis

Afin d'identifier chacune des Q stations de base du réseau, nous leur associons l'indice $q = (1 \dots Q)$. De façon analogue, il est également nécessaire de différencier les antennes d'émission équipant la q^{ieme} station, pour ce faire nous leur associons l'indice $j = (1 \dots t)$. Le couple (q, j) fait alors référence à la j^{ieme} antenne de la q^{ieme} station de base. Finalement, nous identifions chacun des $K_{q,j}$ utilisateurs supportés par l'antenne (q, j) en leur associant l'indice $i = (1 \dots K_{q,j})$. Le triplet (q, j, i) désigne alors le i^{eme} utilisateur supporté par la j^{ieme} antenne de la q^{ieme} station de base. Il est à noter que sous ces conditions nous postulons que le nombre d'utilisateurs supporté par chacune des antennes d'émission du réseau peut être différent.

Le but de notre analyse étant d'évaluer l'impact du multiplexage spatial en liaison descendante, l'ajout d'antennes d'émission supplémentaires au niveau des stations de base est uniquement destiné à multiplier le débit d'information des utilisateurs du réseau mobile par un facteur t . L'antenne (q, j) est ainsi chargée de transmettre $K_{q,j}$ suites de symboles $b_{q,j}^i(m)_{q=1\dots Q, j=1\dots t, i=1\dots K_{q,j}, m \in \mathbf{Z}}$ mutuellement indépendantes dont les éléments sont issus d'une constellation QPSK prenant ainsi les valeurs $\frac{\pm 1 \pm j}{\sqrt{2}}$. De cette façon, chaque station de base est chargée de transmettre t flux indépendants transportant chacun $K_{q,j}$ suites de symboles QPSK, le débit d'information de chacun des $K_{q,j}$ utilisateurs est alors bien multiplié par le nombre d'antennes d'émission t .

Dans la figure 2.2, nous avons représenté la chaîne de multiplexage commune à l'ensemble des stations de base. Nous pouvons observer que les séquences de symboles associées à chacune des antennes d'émission subissent successivement les opérations d'étalement, d'allocation de puissance, d'embrouillage et de mise en forme avant d'être transmises par le réseau d'antenne à travers le canal de propagation. Afin d'introduire notre modèle de signaux, nous allons détailler chacune de ces opérations.

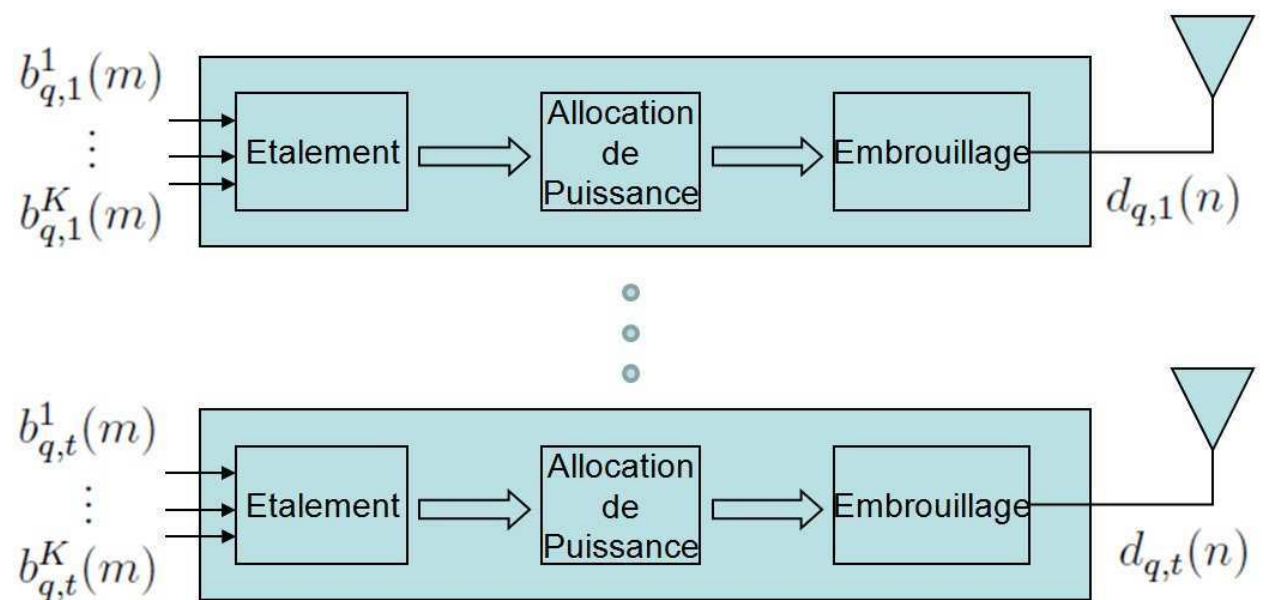


FIGURE 2.2 – Chaîne d'émission

2.1.1 Étalement des données

La séquence de symboles issue de l'antenne d'émission j et destinée à l'utilisateur i est modulée par une séquence de "chips" égaux à $\pm \frac{1}{\sqrt{N}}$, appelée code d'étalement. Plus précisément, à partir de la séquence de symboles $b_{q,j}^i(m)$ et du code d'étalement $c_{q,j}^i(n)_{n=0,\dots,N-1}$, on construit la séquence

$$\dots, c_{q,j}^i(0)b_{q,j}^i(m), \dots, c_{q,j}^i(N-1)b_{q,j}^i(m), c_{q,j}^i(0)b_{q,j}^i(m+1), \dots$$

La taille N du code est appelée facteur d'étalement et caractérise le débit utilisateur. Ainsi, à fréquence "chip" f_c fixée, la fréquence symbole f_s est égale à $\frac{f_c}{N}$ et le débit binaire est égal à $\frac{2f_c}{N}$.

Par souci de simplification des expressions nous faisons le choix d'attribuer un facteur d'étalement N commun à l'ensemble des utilisateurs supportés par la station de base. Dans ce cas de figure, les opérations d'étalement des $K_{q,j}$ séquences de symboles sont synchrones et peuvent s'effectuer en parallèle, ce que nous illustrons par la figure 2.3.

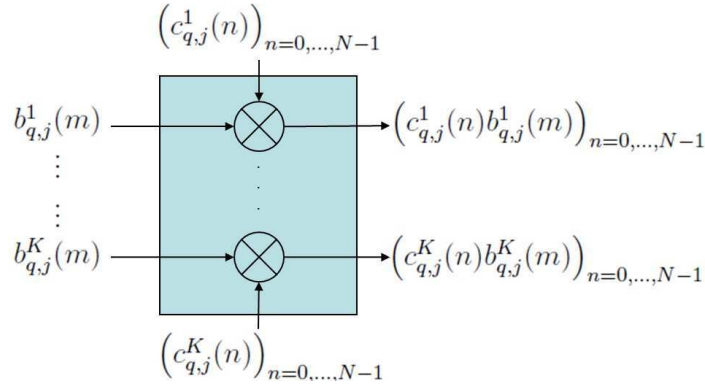


FIGURE 2.3 – Procédure d'étalement

Afin de se rapprocher de la méthode d'accès utilisée dans l'UMTS, les codes d'étalement que nous considérons sont des codes de Walsh-Hadamard orthogonaux vérifiant :

$$\begin{aligned} \sum_{n=0}^{N-1} c_{q,j}^{i_1}(n)c_{q,j}^{i_2}(n) &= 0 \quad \text{si} \quad i_1 \neq i_2 \\ &= 1 \quad \text{sinon} \end{aligned} \quad (2.1)$$

où $c_{q,j}^{i_1}(n)_{n=0,\dots,N-1}$ est le code d'étalement associé à l'utilisateur i_1 supporté par la j^{ime} antenne de la q^{ime} station de base et $c_{q,j}^{i_2}(n)_{n=0,\dots,N-1}$ celui associé à l'utilisateur i_2 .

L'utilisation des codes de Walsh-Hadamard garantissant l'orthogonalité entre les séquences destinées aux différents utilisateurs, il est alors possible de les additionner sans introduire d'interférences. En d'autres termes, à partir de la séquence multi-utilisateurs $\tilde{c}_{q,j}(n)$, définie par

$$\tilde{c}_{q,j}(n) = \sum_{i=1}^{K_{q,j}} c_{q,j}^i(n) b_{q,j}^i, n = 0, \dots, N-1$$

il est possible de retrouver le symbole $b_{q,j}^k$ en procédant à une opération de désétalement de la séquence $\tilde{c}_{q,j}(n)$ à l'aide du code d'étalement $c_{q,j}^k(n)_{n=0,\dots,N-1}$, ce qui se traduit par

$$\sum_{n=0}^{N-1} c_{q,j}^k(n) \tilde{c}_{q,j}(n) = \sum_{i=1}^{K_{q,j}} b_{q,j}^i \sum_{n=0}^{N-1} c_{q,j}^k(n) c_{q,j}^i(n) = b_{q,j}^k$$

2.1.2 Allocation de puissance

Les utilisateurs sont plus ou moins exposés aux phénomènes d'interférences entre cellules et d'atténuation de parcours selon qu'ils se trouvent en bord de cellule ou proche des antennes d'émissions de la station à laquelle ils sont rattachés. Ces contraintes physiques liées à l'environnement de propagation entraînent la nécessité de répartir de façon inégale la puissance totale disponible au niveau de la station de base parmi ses différents utilisateurs afin d'adapter la quantité de puissance nécessaire à l'obtention d'un critère de qualité donné.

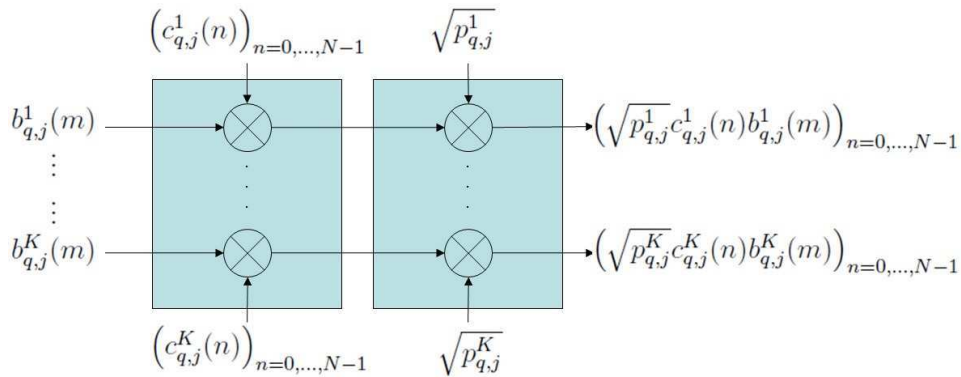


FIGURE 2.4 – Procédure d'allocation de puissance

Au sein de notre modèle, comme illustré par la figure 2.4, nous traduisons ce mécanisme d'allocation et de contrôle de puissance en multipliant les séquences chips

de chaque utilisateur par un facteur d'amplitude $\sqrt{p_{q,j}^i}$ avant de les additionner. Ainsi, l'expression de la séquence chip associée au i^{ime} utilisateur supporté par la j^{ime} antenne de la q^{ime} station de base à l'instant symbole m s'exprime par

$$d_{q,j}^i(mN + n) = \sqrt{p_{q,j}^i} c_{q,j}^i(n) b_{q,j}^i(m)$$

2.1.3 Embrouillage (Scrambling)

Comme illustré par la figure 2.5, la séquence chip résultant de l'addition de chacune des séquences destinées aux différents utilisateurs est ensuite multipliée par une longue séquence de scrambling dont le but principal est de permettre d'identifier et de différencier les différentes antennes émettrices situées dans le voisinage d'un mobile donné. Nous faisons ici le choix d'attribuer des séquences de scrambling différentes à chacune des antennes d'émission d'une même station de base.

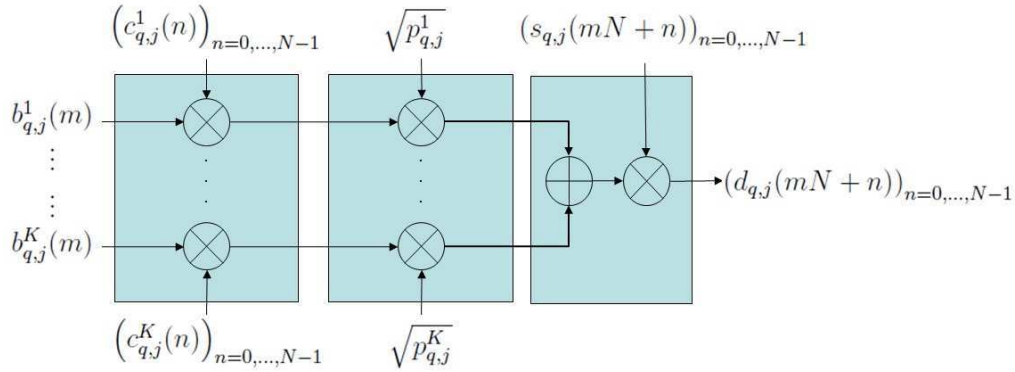


FIGURE 2.5 – Procédure d'embrouillage

Les séquences de scrambling que nous utilisons sont de longues séquences aléatoires indépendantes entre elles dont les entrées i.i.d. prennent les valeurs $\frac{\pm 1 \pm j}{\sqrt{2}}$. En notant $s_{q,j}(n)_{n \in \mathbb{N}}$ le code d'embrouillage associé à la j^{ime} antenne de la q^{ime} station de base, l'expression de la séquence chip multi-utilisateurs émise par celle-ci s'exprime alors par

$$d_{q,j}(mN + n) = s_{q,j}(mN + n) \sum_{i=1}^K \sqrt{p_{q,j}^i} c_{q,j}^i(n) b_{q,j}^i(m)$$

L'action de ces codes de scrambling aléatoires nous permet d'assimiler les signaux émis par le réseau d'antenne associé à la q^{ime} station de base à t séquences aléatoires

$d_{q,j}(n)$ indépendantes entre elles dont la variance $\mu_{q,j}$ est naturellement obtenue à partir de l'expression précédente et vaut

$$\mu_{q,j} = \mathbf{E} [|d_{q,j}(n)|^2] = \frac{\sum_{i=1}^{K_{q,j}} p_{q,j}^i}{N}$$

2.1.4 Mise en forme

Le dernier étage de la chaîne d'émission consiste en une mise en forme de la séquence chip à l'aide d'un filtre d'émission $\tilde{v}(t)$ en racine de Nyquist dont le facteur d'excès de bande est de 0.22 en WCDMA. Le signal à temps continu émis par la j^{ieme} antenne de la q^{ieme} station de base s'écrit ainsi :

$$\tilde{d}_{q,j}(t) = \sum_{n \in \mathbf{Z}} d_{q,j}(n) \tilde{v}(t - nT_c) \quad (2.2)$$

où la durée d'un chip est donnée par $T_c = \frac{1}{f_c}$ (En WCDMA, la fréquence chip vaut 3.84 MHz).

2.2 Modélisation du canal de propagation

La propagation du signal entre la q^{ieme} station de base et le mobile d'intérêt s'effectue dans un environnement susceptible de comporter différents types d'obstacles. Ceci se traduit alors par le fait que le récepteur mobile reçoit plusieurs versions de ce signal (cf. figure 2.6), décalées dans le temps et déphasées l'une par rapport à l'autre. Le jeu de déphasage est susceptible d'entraîner la recombinaison négative de différents trajets synchrones. Ce phénomène connu sous le nom d'évanouissement de Rayleigh entraîne une sélectivité en fréquence du canal de propagation MIMO qu'il est nécessaire de modéliser et de prendre en compte dans notre étude.

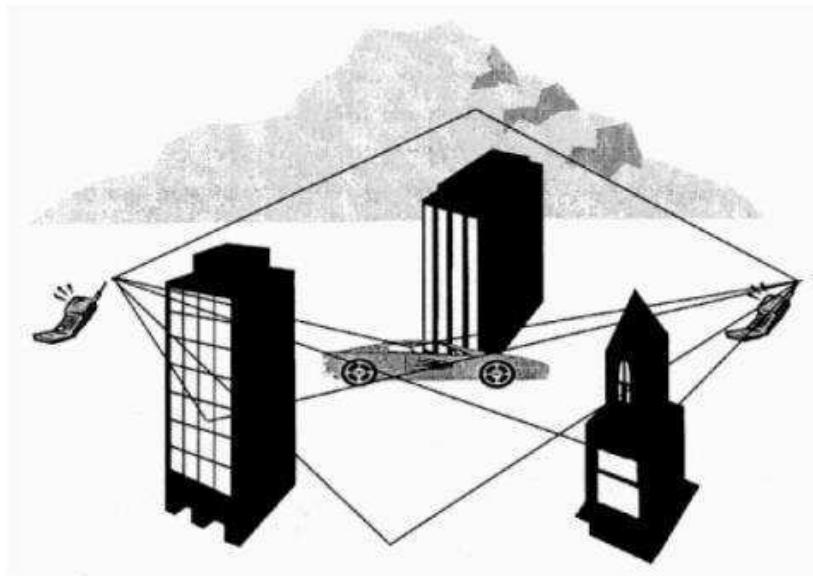


FIGURE 2.6 – Environnement de propagation à trajets multiples

2.2.1 Modèle à temps continu

Nous considérons $\tilde{\mathbf{d}}_{\mathbf{q}}(t)$, un vecteur de dimension $t \times 1$ dont les entrées représentent la valeur à l'instant t du signal émis par chacune des antennes de la q^{ieme} station de base, tel que

$$\tilde{\mathbf{d}}_{\mathbf{q}}(t) = \begin{bmatrix} \tilde{d}_{q,1}(t) \\ \vdots \\ \tilde{d}_{q,t}(t) \end{bmatrix} \quad (2.3)$$

Ces signaux sont alors reçus sur une antenne comportant r capteurs. Nous appelons P le nombre de trajets de propagation. A chacun des trajets p ($p = 1 \dots P$) est associé une atténuation $\sqrt{c_{q,p}}$ et un retard $\tau_{q,p}$ ($\tau_{q,1} < \tau_{q,2} < \dots < \tau_{q,P}$).

Dans ces conditions nous pouvons constituer le vecteur $\tilde{\mathbf{x}}(t)$ de dimension $r \times 1$ par empilement des signaux reçus sur le réseau de capteur

$$\tilde{\mathbf{x}}(t) = \sum_{q=1}^Q \sum_{p=1}^P \sqrt{c_{q,p}} \mathbf{\Lambda}_q^p(t) \tilde{\mathbf{d}}_{\mathbf{q}}(t - \tau_{q,p})$$

où $\mathbf{\Lambda}_q^p(t)$ est une matrice de dimension $r \times t$ dont les entrées représentent les gains complexes du p^{ieme} trajet associés aux différents liens entre le réseau d'antennes d'émission de la q^{ieme} station de base et le réseau d'antennes de réception du mobile d'intérêt.

Il est à noter qu'il existe différentes méthodes visant à générer des matrices $\mathbf{\Lambda}_q^p(t)$ modélisant de façon réaliste un canal de propagation réel, comme par exemple le modèle de Clarke ([23]) qui prend en compte les déphasages géométriques du réseau d'antennes de réception. Toutefois, nous décidons de ne pas recourir à ce type de modélisation fine car le modèle Gaussien que nous présenterons plus en avant dans ce document rend déjà très bien compte du phénomène d'évanouissement de Rayleigh.

2.2.2 Modèle à temps discret

En WCDMA, le signal au rythme chip est conditionné en slot de 2560 chips consécutifs. En pratique la durée correspondante $T_{slot} = 0.67ms$ est négligeable par rapport au temps de cohérence du canal et ce pour des mobilités supérieures à 100 km/h. Il est alors commode de considérer que les matrices $\mathbf{\Lambda}_q^p(t)$ définies précédemment sont constantes sur la durée d'un slot. Cette hypothèse va nous permettre en effet de simplifier grandement notre modèle de propagation à temps discret et de faciliter

les calculs liés à l'étude de performance des récepteurs présentés dans le chapitre suivant.

Conformément à notre hypothèse, nous choisissons donc de découper le signal reçu en "slots" de sorte que l'on peut réexprimer le signal reçu sur le réseau de capteurs à l'aide d'un vecteur de dimension $r \times 1$ de la façon suivante

$$\tilde{\mathbf{x}}_{\mathbf{m}}(t) = \sum_{q=1}^Q \sum_{p=1}^P \sqrt{c_{q,p}} \mathbf{\Lambda}_{q,m}^p \tilde{\mathbf{d}}_{\mathbf{q},\mathbf{m}}(t - \tau_{q,p}) \quad m = 1, \dots, N_{slots} \quad (2.4)$$

où l'on définit par

- N_{slots} , le nombre de slots reçus.
- $\tilde{\mathbf{d}}_{\mathbf{q},\mathbf{m}}(t) = \tilde{\mathbf{d}}_{\mathbf{q}}(t + mT_{slot})$, le signal émis durant le slot m .
- $\tilde{\mathbf{x}}_{\mathbf{m}}(t)$ la portion de signal reçu correspondante.
- $\mathbf{\Lambda}_{q,m}^p = \mathbf{\Lambda}_{q,m}^p(t + mT_{slot})$, une matrice de dimension $r \times t$, constante pendant le slot m , dont les entrées représentent les gains complexes.

Le signal reçu traverse alors un filtre en racine de Nyquist, identique au filtre d'émission $\tilde{v}(t)$, le signal résultant $\mathbf{x}_{\mathbf{m}}(t)$ peut alors s'écrire

$$\mathbf{x}_{\mathbf{m}}(t) = \sum_{q=1}^Q \sum_{p=1}^P \sqrt{c_{q,p}} \mathbf{\Lambda}_{q,m}^p \mathbf{d}_{\mathbf{q},\mathbf{m}}(t - \tau_{q,p}) \quad m = 1, \dots, N_{slots} \quad (2.5)$$

où $\mathbf{d}_{\mathbf{q},\mathbf{m}}(t) = \mathbf{d}_{\mathbf{q}}(t + mT_{slot})$ représente les séquences de chip émises par la q^{ime} station de base et mise en forme par le filtre de Nyquist $v(t)$ obtenu par convolution des filtres d'émission et de réception :

$$\mathbf{d}_{\mathbf{q}}(t) = \sum_{n \in \mathbf{Z}} \mathbf{d}_{\mathbf{q}}(n) v(t - nT_c)$$

Le signal reçu est alors échantillonné à la fréquence $f_e = \frac{1}{T_e} = 2f_c$. A partir du signal à temps continu $\mathbf{x}_{\mathbf{m}}(t)$, nous pouvons former le signal à temps discret $x_m(n) = x_m(nT_e)$ tel que

$$\mathbf{x}_m(n) = \sum_{q=1}^Q \sum_{p=1}^P \sqrt{c_{q,p}} \mathbf{\Lambda}_{q,m}^p \mathbf{d}_{\mathbf{q},\mathbf{m}}(nT_e - \tau_{q,p})$$

En remplaçant alors $\mathbf{d}_{\mathbf{q},\mathbf{m}}(t)$ par son expression nous obtenons

$$\mathbf{x}_m(n) = \sum_{q=1}^Q \sum_{p=1}^P \sqrt{c_{q,p}} \Lambda_{q,m}^p \sum_{n' \in \mathbf{Z}} \mathbf{d}_q(n') v((n - n')T_e - \tau_{q,p}) \quad (2.6)$$

$$= \sum_{q=1}^Q \sum_{l \in \mathbf{Z}} \sum_{p=1}^P \sqrt{c_{q,p}} \Lambda_{q,m}^p v(lT_e - \tau_p) \mathbf{d}_q(n - l) \quad (2.7)$$

On voit alors apparaître la réponse impulsionnelle du canal échantillonnée au rythme T_e , en effet

$$\mathbf{x}_m(n) = \sum_{q=1}^Q \sum_{l \in \mathbf{Z}} \mathbf{H}_{q,m}(l) \mathbf{d}_q(n - l)$$

avec

$$\mathbf{H}_{q,m}(l) = \sum_{p=1}^P \sqrt{c_{q,p}} \Lambda_{q,m}^p v(lT_e - \tau_{q,p})$$

Le filtre de mise en forme $v(t)$ est supposé nul lorsque $|t|$ dépasse une certaine valeur et $(\tau_{q,p} - \tau_{q,1})$ est supposé être négligeable devant la durée d'un slot. Ceci se traduit par le fait que $\mathbf{H}_{q,m}(l)$ est nulle en-deçà et au-delà de certaines valeurs de l . Nous pouvons alors considérer que la réponse impulsionnelle du canal, de taille L , est entièrement comprise entre 0 et $L - 1$.

2.2.3 Modèle aléatoire gaussien

A partir de maintenant nous supprimons l'indice de slot car il ne nous est plus d'aucune utilité dans la suite de nos développements. Nous appelons $\mathbf{h}_{q,j}$ le vecteur de dimension $L \times 1$ composé de la superposition des vecteurs spatiaux de la réponse impulsionnelle échantillonnée correspondant au canal situé entre la j^{ime} antenne de la q^{ime} station de base et le réseau d'antenne de réception

$$\mathbf{h}_{q,j} = \begin{bmatrix} \mathbf{h}_{q,j}(0) \\ \vdots \\ \mathbf{h}_{q,j}(L-1) \end{bmatrix}$$

Nous modéliserons l'ensemble des valeurs prise au cours du temps par $\mathbf{h}_{q,j}$ comme les réalisations d'un vecteur aléatoire gaussien.

Dans la suite de nos développements, il sera utile de comparer de façon juste les performances du filtre de Wiener pour différentes valeurs de t et de L . Pour cette raison nous décidons de normaliser la puissance totale du canal de propagation. De

plus, nous souhaiterions également pouvoir fixer le niveau de puissance reçue au niveau du réseau d'antenne du mobile en provenance des stations de base interférentes relativement au niveau de puissance reçue de référence en provenance de la station de base principale. Pour satisfaire à ces conditions, nous modélisons $\mathbf{H}_{\mathbf{q},\mathbf{m}}(l)$ sous la forme d'une matrice gaussienne de dimension $r \times t$ que nous appelons $\bar{\mathbf{H}}_{\mathbf{q}}^1$, dont les entrées sont des variables aléatoires indépendantes et identiquement distribuées (i.i.d.) centrées de variance $\frac{1}{Lt}$, que l'on multipliera par un facteur d'échelle $\sqrt{a_q}$ (avec $a_1 = 1$). Cette hypothèse est vérifiée si et seulement si :

- les temps de retards $\tau_{q,j}$ sont des multiples de T_c
- les antennes d'émission et de réception sont décorréliées
- les différents trajets physiques ont la même puissance

Remarquons que cette dernière hypothèse, liée au fait que les variances des éléments de $\bar{\mathbf{H}}_{\mathbf{q}}^1$ sont toutes égales à $\frac{1}{Lt}$, n'est pas indispensable car il serait tout à fait possible d'envisager des trajets de puissances différentes dans les calculs présentés un peu plus loin.

Finalement, l'action du canal de propagation situé entre la q^{ime} station de base et l'utilisateur d'intérêt se traduit alors par celle d'un filtre linéaire agissant au rythme chip sur le signal émis, défini par

$$\mathbf{H}_{\mathbf{q}}(z) = \sqrt{a_q} \sum_{l=0}^{L-1} \bar{\mathbf{H}}_{\mathbf{q}}^1 z^{-l} = \sum_{l=0}^{L-1} \mathbf{H}_{\mathbf{q}}^1 z^{-l} \quad (2.8)$$

2.3 Modélisation réaliste du réseau

2.3.1 Géométrie du réseau

Nous souhaitons pouvoir évaluer de la façon la plus réaliste possible les performances du récepteur de Wiener, notamment en prenant en compte dans notre étude la présence d'interférences entre cellules. Pour ce faire, il est nécessaire de fixer la topologie du réseau et de considérer les caractéristiques de l'environnement de propagation de la station de base et du mobile d'intérêt. De la même façon que dans [24], nous approximons la zone de couverture d'une cellule à un disque de rayon $R = 0.557$ kilomètres et nous considérons un agencement hexagonal du réseau, comme l'illustre la figure 2.7.

Dans ces conditions, il est raisonnable de considérer quatre scénarios traduisant les différentes configurations possibles en terme d'interférence inter-cellulaire.

- Scénario 1 : Le mobile se situe à l'intérieur de la cellule et ne reçoit aucune puissance en provenance des cellules environnantes.
- Scénario 2 : Le mobile se situe à l'intérieur de la cellule et reçoit de la puissance en provenance des cellules environnantes. La station de base principale est alors la plus puissante.
- Scénario 3 : Le mobile se situe à l'intersection de deux cellules et reçoit les signaux issus de ces deux stations de base à la même puissance.
- Scénario 4 : Le mobile se situe à l'intersection de trois cellules. Le mobile reçoit les signaux issus de ces trois stations de base à la même puissance.

On remarquera que le scénario 2 constitue le cas d'utilisation rencontré le plus fréquemment alors que les scénarios 3 et 4 correspondent aux cas d'utilisation limite pour lesquels la quantité d'interférences entre cellule est très importante. Enfin le scénario 1 constitue un cas d'étude théorique sans interférences inter cellulaire dont les résultats serviront de référence dans l'étude des trois autres scénarios. Dans chacune de ces configurations spatiales, les performances du récepteur mobile dépendront fortement de sa distance par rapport à chacune des station de base et plus particulièrement de la quantité de puissance reçue en provenance de chacun des émetteurs. Il convient donc d'adopter un modèle de puissance traduisant le comportement réel d'un tel réseau.

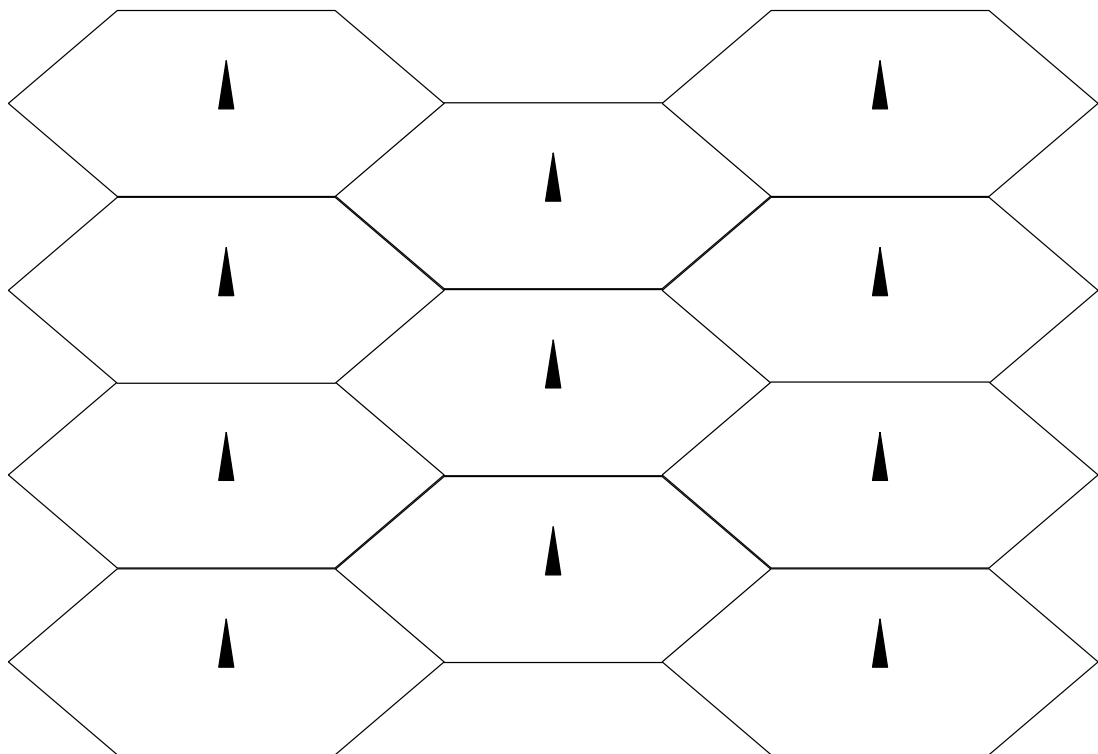


FIGURE 2.7 – Réseau constitué de cellules hexagonales

2.3.2 Niveaux de puissance

Afin que notre modèle des signaux traduise fidèlement la réalité, il est nécessaire de prendre en compte les différents gains et les différentes pertes de puissance susceptibles d'intervenir lors de la transmission. C'est pourquoi nous adoptons le bilan de liaison suivant, inspiré de celui présenté dans [25].

Les différentes stations de bases disposent d'une puissance d'émission maximale $P_{e,max}$ (en dBm) que l'on répartira uniformément parmi ses t antennes d'émission auxquelles nous associons un gain G_{tx} (en dBi). De façon tout à fait standard, nous considérons une perte de puissance L_c (en dB) aux niveau du câblage de la station de base. Sous ces considérations, la puissance irradiée P_i (en dBm) par chacune des antennes est donnée par

$$P_i = P_{e,max} - 10\log_{10}(t) + G_{tx} - L_c$$

Lors de la propagation du signal sur une distance d (en km), la puissance irradiée se voit atténuée par les effets cumulés des pertes de parcours $L_p(d)$ (en dB), des pertes liées à l'effet de masque L_m (en dB) et des pertes d'environnement L_e (en dB). La puissance reçue $P_r(d)$ (en dBm) en provenance d'une station de base est alors donnée par

$$P_r(d) = P_i - L_p(d) - L_m - L_e$$

Il est alors nécessaire de choisir un modèle pour les atténuations de parcours $L_p(d)$ et de fixer chacun des paramètres énoncés ci-dessus.

$P_{e,max}$	43	dBm
G_{tx}	18	dBi
L_c	2	dB
L_m	7	dB
L_e	18	dB

TABLE 2.1 – Paramètres de propagation

Selon la littérature dédiée aux réseaux WCDMA ([26]), les pertes de parcours sont couramment modélisées par

$$L_p(d) = 137.4 + 35.2\log_{10}(d)$$

et les valeurs des différents paramètres sont résumés au sein du tableau 2.1.

2.4 Signal reçu

2.4.1 Expression du signal reçu

Le signal reçu au niveau du réseau d'antennes du mobile d'intérêt est la somme, échantillonnée à la fréquence $f_e = f_c$, des signaux provenant des différentes stations de base à laquelle s'ajoute le bruit thermique que nous considérons comme blanc et gaussien de variance σ^2 . Il peut alors s'exprimer de la façon suivante

$$\mathbf{y}(n) = \sum_{q=1}^Q \sum_{l \in \mathbf{Z}} \mathbf{H}_q^l \mathbf{d}_q(n-l) + \mathbf{v}(n) \quad (2.9)$$

où $\mathbf{v}(n)$ est un vecteur de dimension $r \times 1$ dont les entrées sont des variables aléatoires gaussiennes de variance σ^2 .

Pour rendre notre modèle des signaux cohérent et donner du sens à l'expression du signal reçu dans le modèle de réseau que nous avons adopté, il est nécessaire de relier la variance de bruit σ^2 avec $\left[\frac{E_b}{N_0}\right]_{reel}$ le rapport réel de l'énergie bit reçue par capteur sur le niveau de bruit thermique au sein du récepteur.

Il est classique de considérer que la variance du bruit σ^2 est égale à la puissance de bruit thermique N_0 . Ainsi nous pouvons écrire

$$\sigma^2 = N_0 = E_b \left[\frac{E_b}{N_0}\right]_{reel}^{-1}$$

Pour une modulation MAQ à M points, nous savons que l'énergie bit E_b est égale à l'énergie symbole E_s divisée par $\log_2(M)$. Puisque nous utilisons une modulation QPSK ($M = 4$) l'énergie bit est égale à la moitié de l'énergie symbole. Ensuite, l'étalement de spectre et notamment l'utilisation d'un facteur d'étalement égal à N entraîne que l'énergie symbole E_s est égale à N fois l'énergie chip E_c , ainsi

$$\sigma^2 = \frac{NE_c}{2} \left[\frac{E_b}{N_0}\right]_{reel}^{-1}$$

D'après notre modèle des signaux nous savons que pour une puissance symbole p_s , la puissance chip en provenance de la première station de base et reçue au niveau du réseau d'antennes de l'utilisateur considéré est égale à

$$E_c = \frac{p_s T_c}{tN}$$

Finalement nous obtenons la relation entre la variance de bruit σ^2 de notre modèle et le rapport signal à bruit réel.

$$\sigma^2 = \frac{p_s T_c}{2t} \left[\frac{E_b}{N_0} \right]_{reel}^{-1}$$

Afin de simuler l'effet du bruit thermique, nous voyons qu'il est nécessaire de déterminer le rapport signal à bruit réel, ce que nous montrons dans la section suivante.

2.4.2 Rapport signal à bruit réel

Soit P_e , la puissance (en dBm) allouée par antenne au mobile d'intérêt. Selon le modèle de propagation que nous avons adopté, l'énergie bit utile E_b est donnée par

$$E_b = P_e + G_{tx} - L_c - L_p(d) - L_m - L_e + G_{SF} - 10 \log(B_w)$$

où $G_{SF} = 10 \log_{10}(N)$ est le gain de désétalement (en dB) correspondant au facteur d'étalement N et B_w la bande passante du signal (en Hz). L'ensemble des autres paramètres ont déjà été introduit dans la section précédente.

Dans [25] nous apprenons que le niveau de bruit thermique N_0 au niveau du récepteur en sortie du filtre adapté dépend de la température T (en Kelvin) et du facteur de bruit F (en dB) du récepteur. Il est donc donné par

$$N_0 = 10 \log_{10}(kT) + F$$

ou $k = 1.38065 \times 10^{-23}$ est la constante de Boltzman. En pratique, le facteur de bruit F est compris entre 5 pour les récepteurs de très bonne qualité et 9dB pour les moins bons. Nous fixons ici F à 7dB. Comme notre bande passante est égale à 3.84MHz et que nous fonctionnons à température ambiante (soit $T = 290K$) notre seuil de bruit $N_0 B_w$ est alors constant et égal à $-101dBm$.

Chapitre 3

Démodulation

On se propose maintenant de démoduler les symboles émis par la station de base d'intérêt en direction des mobiles qu'elle supporte. Sans perte de généralité nous nous intéressons uniquement à l'estimation des symboles issus de la séquence $b_{1,1}^1(m)$ destinée à l'utilisateur 1 et transmise par la première antenne d'émission de la première station de base. Nous faisons ici l'hypothèse que la procédure de synchronisation a déjà été réalisée. De cette façon, nous pouvons choisir comme origine des temps l'instant où débute la réception du symbole que l'on souhaite démoduler.

Dans un premier temps, nous allons présenter le récepteur de Wiener optimal. Bien qu'impossible à utiliser en pratique, il est important de le considérer car il nous permet d'introduire les différentes notations qui seront reprises dans le chapitre suivant.

3.1 Le filtre de Wiener optimal

Dans les conditions énoncées précédemment, nous rappelons que l'expression du signal multivarié de dimension r échantillonné au rythme chip reçu au niveau du réseau d'antenne de l'utilisateur d'intérêt à l'instant chip $n, n \in \mathbb{N}$, s'exprime

$$\mathbf{y}(n) = \sum_{q=1}^Q [\mathbf{H}_q(z)] \begin{bmatrix} d_{q,1}(n) \\ \vdots \\ d_{q,t}(n) \end{bmatrix} + \mathbf{v}(n) \quad (3.1)$$

où de façon tout à fait standard, nous considérons la présence d'un bruit blanc additif gaussien centré de variance σ^2 , représenté par le vecteur $\mathbf{v}(n)$ de dimension

r .

Avant d'entrer dans les détails, il est important de rappeler que l'on peut assimiler les codes de scrambling associés à chacune des antennes d'émission aux réalisations successives de variables aléatoires QPSK i.i.d centrées de variance unitaire. Le filtre de Wiener sous-optimal que présentons ici est non causal et à réponse impulsionnelle finie (FIR). L'objectif de ce filtre est d'estimer le symbole émis par la première antenne d'émission de la station de base d'intérêt et destiné à l'utilisateur de référence à partir du signal reçu au niveau du réseau d'antennes de son mobile. Afin d'introduire plus précisément notre filtre d'égalisation, nous allons tout d'abord opérer quelques modifications sur l'expression du signal reçu afin de faire apparaître le symbole $b_{1,1}^1$ que l'on se propose d'estimer.

Nous allons dans un premier temps réécrire le signal reçu en faisant intervenir le modèle polyphase des signaux émis. Ainsi nous regroupons les $2Nr$ premières valeurs du signal observé au sein du vecteur \mathbf{y} de dimension $2Nr$, donné par

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}(0) \\ \vdots \\ \mathbf{y}(N-1) \\ \vdots \\ \mathbf{y}(2N-1) \end{bmatrix} \quad (3.2)$$

Ce vecteur correspond au signal observé pendant la réception de la séquence de chips à laquelle est associée le symbole recherché et celui qui le succède. Afin de modéliser l'action de filtre du canal de propagation sur ces $2N$ échantillons nous introduisons les matrices $\mathcal{H}_{q,j}$ de dimension $2Nr \times 2N + L - 1$ données par

$$\mathcal{H}_{q,j} = \begin{bmatrix} \mathbf{h}_{q,j}(L-1) & \dots & \mathbf{h}_{q,j}(0) & 0 & 0 \\ 0 & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & \mathbf{h}_{q,j}(L-1) & \dots & \mathbf{h}_{q,j}(0) \end{bmatrix} \quad (3.3)$$

Nous introduisons également les vecteurs $\mathbf{d}_{q,j,2N+L-1}$ de dimension $2N + L - 1 \times 1$

donnés par

$$\mathbf{d}_{q,j,2N+L-1} = \begin{bmatrix} d_{q,j}(-L+1) \\ \vdots \\ d_{q,j}(0) \\ \vdots \\ d_{q,j}(N-1) \\ \vdots \\ d_{q,j}(2N-1) \end{bmatrix}$$

Nous pouvons alors réécrire le signal reçu sous la forme du modèle vectoriel suivant :

$$\mathbf{y}_{2N} = \mathcal{H}_{1,1}\mathbf{d}_{1,1,2N+L-1} + \sum_{j \neq 1} \mathcal{H}_{1,j}\mathbf{d}_{1,j,2N+L-1} + \sum_{q \neq 1} \sum_{j=1}^t \mathcal{H}_{q,j}\mathbf{d}_{q,j,2N+L-1} + \mathbf{v}_{2N}$$

où

$$\mathbf{v}_{2N} = \begin{bmatrix} \mathbf{v}(0) \\ \vdots \\ \mathbf{v}(N-1) \\ \vdots \\ \mathbf{v}(2N-1) \end{bmatrix} \quad (3.4)$$

Nous devons maintenant faire apparaître le symbole $b_{1,1}^1(0)$ à partir de l'expression précédente. Pour ce faire, nous allons découper le vecteur $\mathbf{d}_{1,1,2N+L-1}$ en 3 parties, ainsi nous distinguerons les parties correspondant au symbole courant $b_{1,1}^1(0)$, à $b_{1,1}^1(-1)$ qui le précède et à $b_{1,1}^1(1)$ qui le succède. Nous définissons ainsi les vecteurs de dimension $N \times 1$

$$\begin{aligned} \mathbf{d}_N^{(0)} &= \begin{bmatrix} d(0) \\ \vdots \\ d(N-1) \end{bmatrix} \\ \mathbf{d}_N^{(-1)} &= \begin{bmatrix} d(-N) \\ \vdots \\ d(-1) \end{bmatrix} \\ \mathbf{d}_N^{(1)} &= \begin{bmatrix} d(N) \\ \vdots \\ d(2N-1) \end{bmatrix} \end{aligned}$$

Nous introduisons également les matrices de dimension $2Nr \times N$ traduisant l'action de filtre du canal de propagation sur les séquences $\mathbf{d}_N^{(0)}$, $\mathbf{d}_N^{(1)}$ et $\mathbf{d}_N^{(-1)}$

$$\begin{aligned}\mathcal{H}^{(0)} &= \begin{bmatrix} \mathbf{H}^{(0)} \\ \mathbf{H}^{(1)} \end{bmatrix} \\ \mathcal{H}^{(1)} &= \begin{bmatrix} 0 \\ \mathbf{H}^{(0)} \end{bmatrix} \\ \mathcal{H}^{(-1)} &= \begin{bmatrix} \mathbf{H}^{(1)} \\ 0 \end{bmatrix}\end{aligned}$$

Les blocs de dimension $N \times N$ qui composent $\mathcal{H}^{(0)}$, $\mathcal{H}^{(1)}$ et $\mathcal{H}^{(-1)}$ sont eux-mêmes définis par

$$\begin{aligned}\mathbf{H}^{(0)} &= \begin{bmatrix} \mathbf{h}_{1,1}(0) & 0 & & 0 \\ \vdots & \ddots & & \\ \mathbf{h}_{1,1}(L-1) & & & \\ 0 & \ddots & \ddots & \\ & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & \mathbf{h}_{1,1}(L-1) & \dots & \mathbf{h}_{1,1}(0) \end{bmatrix} \\ \mathbf{H}^{(1)} &= \begin{bmatrix} 0 & 0 & \mathbf{h}_{1,1}(L-1) & \dots & \mathbf{h}_{1,1}(1) \\ & & & \ddots & \vdots \\ & & & & \mathbf{h}_{1,1}(L-1) \\ & & & & 0 \\ 0 & & & & 0 \end{bmatrix}\end{aligned}$$

A l'aide de ces notations, le premier terme de l'expression précédente devient

$$\mathcal{H}_{1,1} \mathbf{d}_{2N+L-1} = \mathcal{H}^{(0)} \mathbf{d}_N^{(0)} + \mathcal{H}^{(-1)} \mathbf{d}_N^{(-1)} + \mathcal{H}^{(1)} \mathbf{d}_N^{(1)}$$

Il nous reste maintenant à décomposer les vecteurs $\mathbf{d}_N^{(0)}$, $\mathbf{d}_N^{(-1)}$ et $\mathbf{d}_N^{(1)}$ afin de faire apparaître le code de scrambling de la première antenne de la station de base utile, les codes d'étalement de ses utilisateurs, leurs énergies et les symboles émis. Pour ce faire, nous définissons les matrices $N \times N$ contenant les codes de scrambling associés aux séquences de chips précédemment définies

$$\mathbf{S}^{(0)} = \begin{bmatrix} s_{1,1}(0) & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & s_{1,1}(N-1) \end{bmatrix}$$

$$\mathbf{S}^{(1)} = \begin{bmatrix} s_{1,1}(N) & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & s_{1,1}(2N-1) \end{bmatrix}$$

$$\mathbf{S}^{(-1)} = \begin{bmatrix} s_{1,1}(-N) & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & s_{1,1}(-1) \end{bmatrix}$$

Nous introduisons également la matrice de dimension $N \times K_{1,1}$ contenant les codes d'étalement associés aux séquences de chips contenues dans les vecteurs $\mathbf{d}_N^{(0)}$, $\mathbf{d}_N^{(-1)}$ et $\mathbf{d}_N^{(1)}$

$$\mathbf{C} = \begin{bmatrix} c_{1,1}^1(0) & \dots & c_{1,1}^{K_{1,1}}(0) \\ \vdots & \ddots & \vdots \\ c_{1,1}^1(N-1) & \dots & c_{1,1}^{K_{1,1}}(N-1) \end{bmatrix}$$

Nous traduisons l'allocation de puissance par utilisateur à l'aide d'une matrice $K_{1,1} \times K_{1,1}$

$$\mathbf{P} = \begin{bmatrix} p_{1,1}^1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & p_{1,1}^{K_{1,1}} \end{bmatrix}$$

Et nous rassemblons les symboles destinés aux $K_{1,1}$ utilisateurs aux différents instants symboles dans les vecteurs de dimension $K_{1,1} \times 1$

$$\mathbf{b}_{1,1}^{(0)} = \begin{bmatrix} b_{1,1}^1(0) \\ \vdots \\ b_{1,1}^{K_{1,1}}(0) \end{bmatrix}$$

$$\mathbf{b}_{1,1}^{(-1)} = \begin{bmatrix} b_{1,1}^1(-1) \\ \vdots \\ b_{1,1}^{K_{1,1}}(-1) \end{bmatrix}$$

$$\mathbf{b}_{1,1}^{(1)} = \begin{bmatrix} b_{1,1}^1(1) \\ \vdots \\ b_{1,1}^{K_{1,1}}(1) \end{bmatrix}$$

Ces notations nous permettent de décomposer les vecteurs $\mathbf{d}_N^{(0)}$, $\mathbf{d}_N^{(-1)}$ et $\mathbf{d}_N^{(1)}$ de la façon suivante

$$\begin{aligned} \mathbf{d}_N^{(0)} &= \mathbf{S}^{(0)} \mathbf{C} \sqrt{\mathbf{P}} \mathbf{b}_{1,1}^{(0)} \\ \mathbf{d}_N^{(1)} &= \mathbf{S}^{(1)} \mathbf{C} \sqrt{\mathbf{P}} \mathbf{b}_{1,1}^{(1)} \\ \mathbf{d}_N^{(-1)} &= \mathbf{S}^{(-1)} \mathbf{C} \sqrt{\mathbf{P}} \mathbf{b}_{1,1}^{(-1)} \end{aligned}$$

L'expression vectorielle du signal reçu peut alors être écrite de la façon suivante

$$\begin{aligned} \mathbf{y}_{2N} = & \mathcal{H}^{(0)} \mathbf{S}^{(0)} \mathbf{C} \sqrt{\mathbf{P}} \mathbf{b}_{1,1}^{(0)} & + \\ & \mathcal{H}^{(-1)} \mathbf{S}^{(1)} \mathbf{C} \sqrt{\mathbf{P}} \mathbf{b}_{1,1}^{(1)} & + \\ & \mathcal{H}^{(1)} \mathbf{S}^{(-1)} \mathbf{C} \sqrt{\mathbf{P}} \mathbf{b}_{1,1}^{(-1)} & + \\ & \sum_{j \neq 1} \mathcal{H}_{1,j} \mathbf{d}_{1,j,2N+L-1} & + \\ & \sum_{q \neq 1} \sum_{j=1}^t \mathcal{H}_{q,j} \mathbf{d}_{q,j,2N+L-1} + \mathbf{v}_{2N} \end{aligned}$$

Le filtre de Wiener optimal $\mathbf{g}_{2N,opt}$ est celui qui minimise l'espérance de l'erreur quadratique moyenne entre sa sortie et le symbole recherché

$$\mathbf{g}_{2N,opt} = \underset{\mathbf{g}}{\operatorname{argmin}} \mathbf{E} \left[|\mathbf{g}^H \mathbf{y}_{2N} - b_{1,1}^1(0)|^2 \right]$$

L'expression de ce filtre est donnée par

$$\mathbf{g}_{2N,opt} = \mathbf{E} [\mathbf{y}_{2N} \mathbf{y}_{2N}^H]^{-1} \mathbf{E} [\mathbf{y}_{2N} (b_{1,1}^1(0))^*]$$

En remplaçant \mathbf{y}_{2N} par son expression, nous obtenons après calcul

$$\mathbf{g}_{2N,opt} = \mathcal{R}^{-1} \mathcal{H}_{1,1}^{(0)} \mathbf{S}^{(0)} \mathbf{c}_{N,1} \sqrt{p_{1,1}^1}$$

où $\mathbf{c}_{N,1}$ est la première colonne de la matrice de code \mathbf{C} contenant le code d'étalement du canal utile, $p_{1,1}^1$ est la puissance allouée au premier utilisateur supporté par la première antenne d'émission de la première station de base et la matrice \mathcal{R} est donnée par

$$\begin{aligned} \mathbf{R} = & \sum_{l=-1}^1 \mathcal{H}^{(l)} \mathbf{S}^{(l)} \mathbf{C} \mathbf{P} \mathbf{C}^H (\mathbf{S}^{(l)})^H (\mathcal{H}^{(l)})^H + \\ & \sum_{j \neq 1} \mu_{1,j} \mathcal{H}_{1,j} (\mathcal{H}_{1,j})^H + \\ & \sum_{q \neq 1} \sum_{j=1}^t \mu_{q,j} \mathcal{H}_{q,j} \mathcal{H}_{q,j}^H + \\ & \sigma^2 \mathbf{I}_{2N} \end{aligned}$$

Dans l'expression ci-dessus, il est important de remarquer que nous faisons l'hypothèse que les signaux transmis par les antennes interférentes associées à la première station de base, ainsi que ceux transmis par les antennes des stations de base interférentes, sont des séquences de chips indépendantes entre elles i.i.d. stationnaires de variance égale à $\mu_{q,j}$ pour $q = 1, j \neq 1$ et $q \neq 1, j = 1, \dots, t$. Nous contournons de cette façon l'impossibilité connaître précisément au niveau du récepteur les séquences de scrambling et les codes d'étalements associés à l'ensemble des antennes interférentes.

Ce récepteur nous permet alors d'estimer le symbole utile de la façon suivante :

$$b_{1,1}^1(0) = \mathbf{g}_{2N,opt}^H \mathbf{y}_{2N}$$

Ce démodulateur ne peut être mis en oeuvre en pratique que si l'on connaît les canaux de propagation et les puissances de toutes les stations de base interférentes, ainsi que les codes d'étalement et les énergies symbole de tous les utilisateurs de la station utile, ce qui est rarement le cas en pratique. Il est donc nécessaire de se tourner vers un démodulateur sous-optimal mais néanmoins implémentable.

3.2 Le filtre de Wiener sous-optimal

La structure sous-optimale que nous allons présenter dans ce paragraphe est une généralisation au cas MIMO de celle présentée par [27] et [28]. Elle est composée d'un égaliseur au rythme chip, suivi d'un étage de désembrouillage et de désétalement. L'égaliseur au rythme chip permet de produire un estimé de la séquence de chip transportant le symbole à démoduler. En notant $\mathbf{g}(z)$ le filtre égaliseur, l'expression du chip estimé à l'instant n devient

$$\hat{d}(n) = [\mathbf{g}(z)] \mathbf{y}(n)$$

Après filtrage, il devient possible de procéder aux opérations de désembrouillage et de désétalement afin de produire un estimé du symbole recherché. Pour ce faire, nous considérons la concaténation de N estimés successifs au sein du vecteur $\hat{\mathbf{d}}$ défini par

$$\hat{\mathbf{d}} = \begin{bmatrix} \hat{d}(0) \\ \vdots \\ \hat{d}(N-1) \end{bmatrix}$$

L'estimé du symbole recherché est alors donné par

$$\hat{b}_{1,1}^1(0) = \mathbf{c}_{1,N}^H \mathbf{S}_{1,1}^H \hat{\mathbf{d}}_N \quad (3.5)$$

où \mathbf{c}_1 est la première colonne de la matrice de code \mathbf{C} .

Le filtre de Wiener sous-optimal que nous considérons ici est donné par

$$\mathbf{g}(z) = \sum_{l=-(N-1)}^N \mathbf{g}_l z^{-l}$$

Il nous permet de définir un vecteur de dimension $2Nr$ résultant de la concaténation des coefficients du filtre considéré ci-dessus, donné par

$$\mathbf{g} = [\mathbf{g}_N, \dots, \mathbf{g}_{-(N-1)}]$$

Le filtre de Wiener sous-optimal constitue une alternative implémentable au filtre de Wiener optimal présenté précédemment car il requiert uniquement la connaissance des canaux utiles et de la variance des séquences transmises par chacune des stations de base. Il est construit sous l'hypothèse que le signal reçu est stationnaire au niveau chip et que les séquences de chip $(d_{q,j}(n))_{n \in \mathbb{Z}}$ transmises par chacune des antennes d'émission du réseau sont des suites i.i.d. décorréées de variance $\mu_{q,j}$. A la différence du filtre de Wiener optimal, qui minimise l'erreur quadratique moyenne entre sa sortie et le symbole à estimer, le filtre de Wiener sous-optimal minimise l'erreur quadratique moyenne entre sa sortie et la séquence $(d_{1,1}(n))_{n \in \mathbb{Z}}$, ce qui se traduit par

$$\mathbf{g}_{2N,MMSE} = \underset{\mathbf{g}}{\operatorname{argmin}} \mathbf{E} \left[|\mathbf{g}^H \mathbf{y}_{2N} - d_{1,1}(n)|^2 \right]$$

Le filtre est alors donné par la formule

$$\mathbf{g}_{2N,MMSE} = \left(\sum_{q=1}^Q \sum_{j=1}^t \mu_{q,j} \mathcal{H}_{q,j} \mathcal{H}_{q,j}^H + \sigma^2 \mathbf{I}_r \right)^{-1} \mathbf{h}_{1,1} \quad (3.6)$$

où $\mathbf{h}_{1,1}$ est un vecteur de dimension $2Nr$, dont les Nr premier éléments sont nuls, donné par

$$\mathbf{h}_{1,1} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \mathbf{h}_{1,1}(0) \\ \dots \\ \mathbf{h}_{1,1}(L-1) \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (3.7)$$

Chapitre 4

Analyse de performance

4.1 Indice de performance

L'indicateur de performance que nous avons retenu dans notre étude est le taux d'erreur binaire (TEB) en sortie du démodulateur. Bien qu'il ne soit pas possible de calculer explicitement ce TEB il est toutefois possible d'en obtenir une bonne approximation à l'aide du rapport signal à bruit plus interférences (SINR) en sortie du récepteur en généralisant au cas MIMO les résultats de [29] et [30]. De cette façon, nous pouvons considérer que l'ensemble des interférences et du bruit en sortie du récepteur MMSE est gaussien centré et uniquement caractérisé par sa variance lorsque le facteur d'étalement N , le nombre d'utilisateurs $K_{1,1}$ et les nombres d'antennes t en émission et r en réception sont suffisamment élevés. Dans ces conditions, si nous appelons β_N la valeur réelle du SINR en sortie du récepteur de Wiener, le taux d'erreur binaire associé à la modulation QPSK peut alors être calculé à l'aide de la formule suivante :

$$TEB = Q\left(\sqrt{\beta_N}\right)$$

ou $Q(x)$ est la fonction queue de gaussienne définie par

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{+\infty} e^{-\frac{t^2}{2}} dt$$

Cependant, ce rapport signal à bruit réel dépend des codes d'étalement et de scrambling utilisés ainsi que des réalisations du canal de propagation et des énergies attribuées aux différents utilisateurs au niveau du réseau d'antennes d'émission de chacune des stations de base, il est de ce fait difficile à interpréter.

Afin de s'affranchir de ces dépendances il est alors classique d'étudier la moyenne statistique de ce TEB. Or, les réalisations du canal de propagation et des codes de scrambling entraînent la variation du SINR d'un symbole à l'autre. Dans cette partie, nous nous attacherons donc à démontrer que cette variation du TEB autour de sa valeur moyenne est d'autant plus faible que

- les nombres d'utilisateurs $K_{q,j}$ et le facteur d'étalement N sont grands
- les nombres d'antennes t en émission et r en réception sont grands

Pour ce faire, nous allons dans un premier temps étudier le comportement de β_N lorsque les nombre d'utilisateurs $K_{q,j}$ et le facteur d'étalement N tendent vers l'infini en conservant un rapport $\alpha_{q,j} = \frac{K_{q,j}}{N}$ constant. Dans ces conditions, nous allons montrer que β_N converge presque sûrement vers une limite β que nous allons calculer. Nous généraliserons ainsi les résultats de ([1]) au cas MIMO dans un environnement de propagation sélectif en fréquence en ayant recours à des codes d'étalement orthogonaux et une distribution non uniforme des énergies symboles. Ensuite, nous étudierons le comportement de β lorsque les nombres d'antennes t en émission et r en réception tendent vers l'infini en conservant un rapport $\frac{r}{t}$ constant. Nous démontrerons ainsi que β converge presque sûrement vers β^* , dont l'expression est indépendante des valeurs particulières prises par les coefficients de l'équivalent à temps discret du canal de propagation.

4.2 Calcul du SINR

Afin d'évaluer les performances du récepteur associé au lien décrit précédemment, nous devons développer l'équation (3.5) pour différencier le signal utile provenant de l'antenne d'intérêt et les signaux d'interférences dûs aux autres antennes et stations de base. Dans un premier temps, il est nécessaire de traduire les différentes actions de filtre sous forme matricielle. Pour cette raison, nous récrivons l'expression de $\hat{d}(n)$ de la façon suivante

$$\hat{d}(n) = \sum_{q=1}^Q \sum_{j=1}^t [\mathbf{g}(z) \mathbf{h}_{q,j}(z)] d_{q,j}(n) + [\mathbf{g}(z)] v(n) \quad (4.1)$$

$$= \sum_{q=1}^Q \sum_{j=1}^t [\mathbf{f}_{q,j}(z)] d_{q,j}(n) + [\mathbf{g}(z)] v(n) \quad (4.2)$$

où nous décidons de nommer par $\mathbf{f}_{q,j}(z)$ l'action conjuguée du filtre égaliseur $\mathbf{g}(z)$ et de $\mathbf{h}_{q,j}(z)$, la j^{ieme} colonne de la matrice $\mathbf{H}_q(z)$ du canal de propagation telle qu'elle

est décrite par (2.8), ainsi

$$f_{q,j}(z) = \sum_{l=-(N-1)}^{N+L} f_{q,j}(l)^k z^{-l} = \mathbf{g}(z) \mathbf{h}_{q,j}(z)$$

Afin de collecter l'ensemble de l'énergie correspondant au symbole reçu à l'instant 0, il est nécessaire d'appliquer le filtre $\mathbf{f}_{q,j}(z)$ à la séquence chip correspondant aux symboles émis aux instants $-2, -1, 0, 1$ par la j^{ime} antenne de la q^{ime} station de base. Pour ce faire, nous introduisons le vecteur de dimension $4N$ contenant les chips correspondant à ces symboles, donné par

$$\mathbf{d}_{q,j,4N} = \begin{bmatrix} \hat{d}_{q,j}(-2N) \\ \vdots \\ \hat{d}_{q,j}(-N-1) \\ \hat{d}_{q,j}(-N) \\ \vdots \\ \hat{d}_{q,j}(-1) \\ \hat{d}_{q,j}(0) \\ \vdots \\ \hat{d}_{q,j}(N-1) \\ \hat{d}_{q,j}(N) \\ \vdots \\ \hat{d}_{q,j}(2N-1) \end{bmatrix}$$

Nous introduisons $\mathcal{F}_{q,j}$, une matrice de Sylvester de dimension $N \times 4N$ définie par

$$\mathcal{F}_{q,j} = \mathcal{G} \mathcal{H}_{q,j}^{3N} \quad (4.3)$$

où $\mathcal{H}_{q,j}^{3N}$ est une matrice de Sylvester de dimension $3N \times 4N$ traduisant l'action de filtre du canal de propagation, définie par

$$\mathcal{H}_{q,j}^{3N} = \begin{bmatrix} \mathbf{H}_{q,j}^1 & \mathbf{H}_{q,j}^0 & 0 & 0 \\ 0 & \mathbf{H}_{q,j}^1 & \mathbf{H}_{q,j}^0 & 0 \\ 0 & 0 & \mathbf{H}_{q,j}^1 & \mathbf{H}_{q,j}^0 \end{bmatrix}$$

et où \mathcal{G} , une matrice de Sylvester de dimension $N \times 3N$ traduisant l'action du filtre

MMSE, définie par

$$\mathcal{G} = \begin{bmatrix} \mathbf{g}_N & \cdots & \mathbf{g}_{-(N-1)} & 0 & & 0 \\ 0 & \mathbf{g}_N & \cdots & \mathbf{g}_{-(N-1)} & & \\ & & \ddots & \ddots & \ddots & \\ & & & \mathbf{g}_N & \cdots & \mathbf{g}_{-(N-1)} & 0 \\ 0 & & & 0 & \mathbf{g}_N & \cdots & \mathbf{g}_{-(N-1)} \end{bmatrix} \quad (4.4)$$

On peut alors exprimer $\hat{\mathbf{d}}_N$ de la façon suivante

$$\hat{\mathbf{d}}_N = \sum_{q=1}^Q \sum_{j=1}^t \mathcal{F}_{q,j} \mathbf{d}_{q,j,4N} + \mathcal{G} \mathbf{v}_{3N}$$

où \mathbf{v}_{3N} est un vecteur donné par

$$\mathbf{v}_{3N} = \begin{bmatrix} v(-N) \\ \vdots \\ v(-1) \\ v(0) \\ \vdots \\ v(N-1) \\ v(N) \\ \vdots \\ v(2N-1) \end{bmatrix}$$

dont les entrées sont des variables aléatoires gaussiennes i.i.d. centrées de variance σ^2 .

La poursuite de notre étude nécessite de découper les vecteurs $\mathbf{d}_{q,j,4N}$ en 4 parties. Nous introduisons ainsi les vecteurs de dimension $N \times 1$ suivant

$$\mathbf{d}_{q,j,N}^{(-2)} = \begin{bmatrix} d_{q,j}(-2N) \\ \vdots \\ d_{q,j}(-N-1) \end{bmatrix}$$

$$\mathbf{d}_{q,j,N}^{(-1)} = \begin{bmatrix} d_{q,j}(-N) \\ \vdots \\ d_{q,j}(-1) \end{bmatrix}$$

$$\mathbf{d}_{q,j,N}^{(0)} = \begin{bmatrix} d_{q,j}(0) \\ \vdots \\ d_{q,j}(N-1) \end{bmatrix}$$

$$\mathbf{d}_{q,j,N}^{(1)} = \begin{bmatrix} d_{q,j}(N) \\ \vdots \\ d_{q,j}(2N-1) \end{bmatrix}$$

Également, nous partitionnons $\mathcal{F}_{q,j}$ en 4 blocs de dimension $N \times N$ de sorte que

$$\mathcal{F}_{q,j} = \left[F_{q,j}^{(-2)}, F_{q,j}^{(-1)}, F_{q,j}^{(0)}, F_{q,j}^{(1)} \right]$$

Nous pouvons alors réécrire $\hat{\mathbf{d}}_N$ de la façon suivante

$$\hat{\mathbf{d}}_N = \sum_{q=1}^Q \sum_{j=1}^t \sum_{l=-2}^1 F_{q,j}^{(l)} \mathbf{d}_{q,j,N}^{(l)} + \mathcal{G} \mathbf{v}_{3N} \quad (4.5)$$

Nous décomposons alors les vecteurs $\mathbf{d}_{q,j,N}^{(-2)}$, $\mathbf{d}_{q,j,N}^{(-1)}$ et $\mathbf{d}_{q,j,N}^{(1)}$ de la façon suivante

$$\begin{aligned} \mathbf{d}_{q,j,N}^{(-2)} &= \mathbf{S}_{q,j}^{(-2)} \mathbf{C}_{q,j} \sqrt{\mathbf{P}_{q,j}} \mathbf{b}_{q,j}^{(-2)} \\ \mathbf{d}_{q,j,N}^{(-1)} &= \mathbf{S}_{q,j}^{(-1)} \mathbf{C}_{q,j} \sqrt{\mathbf{P}_{q,j}} \mathbf{b}_{q,j}^{(-1)} \\ \mathbf{d}_{q,j,N}^{(1)} &= \mathbf{S}_{q,j}^{(1)} \mathbf{C}_{q,j} \sqrt{\mathbf{P}_{q,j}} \mathbf{b}_{q,j}^{(1)} \end{aligned}$$

Afin de faire apparaître explicitement les contributions du signal utile et de l'interférence multi utilisateur, il est nécessaire de réécrire $\mathbf{d}_{1,1,N}^{(0)}$ de la façon suivante

$$\begin{aligned} \mathbf{d}_{1,1,N}^{(0)} &= \mathbf{S}_{1,1}^{(0)} \mathbf{c}_{1,1} \sqrt{p_{1,1}} b_{1,1}^1(0) + \\ &\quad \mathbf{S}_{1,1}^{(0)} \mathbf{C}_{1,1}^2 \sqrt{\mathbf{P}_{1,1}^2} \mathbf{b}_{1,1,2}^{(0)} \end{aligned}$$

où nous introduisons le vecteur de dimension $(K_{1,1} - 1) \times 1$

$$\mathbf{b}_{1,1,2}^{(0)} = \begin{bmatrix} b_{1,1}^2(0) \\ \vdots \\ b_{1,1}^{K_{1,1}}(0) \end{bmatrix}$$

contenant les symboles émis par la première antenne de la station utile, excepté le symbole recherché, la matrice de dimension $N \times (K_{1,1} - 1)$

$$\mathbf{C}_{1,1}^2 = \begin{bmatrix} c_{1,1}^2(0) & \dots & c_{1,1}^{K_{1,1}}(0) \\ \vdots & \ddots & \vdots \\ c_{1,1}^2(N-1) & \dots & c_{1,1}^{K_{1,1}}(N-1) \end{bmatrix}$$

contenant les codes d'étalement associés aux symboles contenus dans $\mathbf{b}_{1,1,2}^{(0)}$, qui est tout simplement la matrice de code de la première antenne privée de sa première colonne. Nous introduisons également une matrice diagonale de dimension $(K_{1,1} - 1) \times (K_{1,1} - 1)$

$$\mathbf{P}_{1,1}^2 = \begin{bmatrix} p_{1,1}^2 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & p_{1,1}^{K_{1,1}} \end{bmatrix}$$

contenant les valeurs de puissance allouée à ces mêmes symboles.

Toutes ces manipulations nous permettent alors de récrire le vecteur \hat{d}_N de la façon suivante

$$\begin{aligned} \hat{\mathbf{d}}_N = & \mathbf{F}_{1,1}^{(0)} \mathbf{S}_{1,1}^{(0)} \mathbf{c}_{1,1} \sqrt{p_{1,1}^1} b_{1,1}^1(0) + \\ & \mathbf{F}_{1,1}^{(0)} \mathbf{S}_{1,1}^{(0)} \mathbf{C}_{1,1}^2 \sqrt{\mathbf{P}_{1,1}^2} \mathbf{b}_{1,1,2}^{(0)} + \\ & \sum_{l \neq 0} \mathbf{F}_{1,1}^{(l)} \mathbf{S}_{1,1}^{(l)} \mathbf{c}_{1,1} \sqrt{\mathbf{P}_{1,1}} \mathbf{b}_{1,1}^{(l)} + \\ & \sum_{j=2}^t \sum_{l=-2}^1 \mathbf{F}_{1,j}^{(l)} \mathbf{S}_{1,j}^{(l)} \mathbf{c}_{1,j} \sqrt{\mathbf{P}_{1,j}} \mathbf{b}_{1,j}^{(l)} + \\ & \sum_{q=2}^Q \sum_{j=1}^t \sum_{l=-2}^1 \mathbf{F}_{q,j}^{(l)} \mathbf{S}_{q,j}^{(l)} \mathbf{c}_{q,j} \sqrt{\mathbf{P}_{q,j}} \mathbf{b}_{q,j}^{(l)} + \\ & \mathcal{G}\mathcal{V}^{3N} \end{aligned}$$

Ensuite, en procédant aux opérations de désembrouillage et de désétalement, nous obtenons l'expression du symbole estimé faisant intervenir de façon explicite les contributions du signal utile, des interférences de nature différentes, et de la remonté de bruit inhérente au filtrage.

$$\begin{aligned}
\hat{b}_{1,1}^1(0) &= \mathbf{c}_{1,1}^H \left(\mathbf{S}_{1,1}^{(0)} \right)^H \mathbf{F}_{1,1}^0 \mathbf{S}_{1,1}^{(0)} \mathbf{c}_{1,1} \sqrt{p_{1,1}^1} b_{1,1}^1(0) + \\
&\quad \mathbf{c}_{1,1}^H \left(\mathbf{S}_{1,1}^{(0)} \right)^H \mathbf{F}_{1,1}^{(0)} \mathbf{S}_{1,1}^{(0)} \mathbf{C}_{1,1}^2 \sqrt{\mathbf{P}_{1,1}^2} \mathbf{b}_{1,1,2}^{(0)} + \\
&\quad \sum_{l \neq 0} \mathbf{c}_{1,1}^H \left(\mathbf{S}_{1,1}^{(0)} \right)^H \mathbf{F}_{1,1}^l \mathbf{S}_{1,1}^{(l)} \mathbf{C}_{1,1} \sqrt{\mathbf{P}_{1,1}} \mathbf{b}_{1,1}^{(l)} + \\
&\quad \sum_{j=2}^t \sum_{l=-2}^1 \mathbf{c}_{1,1}^H \left(\mathbf{S}_{1,1}^{(0)} \right)^H \mathbf{F}_{1,j}^{(l)} \mathbf{S}_{1,j}^{(l)} \mathbf{C}_{1,j} \sqrt{\mathbf{P}_{1,j}} \mathbf{b}_{1,j}^{(l)} + \\
&\quad \sum_{q=2}^Q \sum_{j=1}^t \sum_{l=-2}^1 \mathbf{c}_{1,1}^H \left(\mathbf{S}_{1,1}^{(0)} \right)^H \mathbf{F}_{q,j}^{(l)} \mathbf{S}_{q,j}^{(l)} \mathbf{C}_{q,j} \sqrt{\mathbf{P}_{q,j}} \mathbf{b}_{q,j}^{(l)} + \\
&\quad \mathbf{c}_{1,1}^H \left(\mathbf{S}_{1,1}^{(0)} \right)^H \mathcal{G} \mathcal{V}^{3N}
\end{aligned}$$

Finalement, nous aboutissons à l'expression de β_N , le SINR associé à \hat{b} , dont la dépendance temporelle est liée à la présence des séquences de scrambling différentes à chaque instant symbole.

$$\beta_N = \frac{A}{B + \sum_{l \neq 0} T_l + \sum_{j=2}^t \sum_{l=-1}^2 U_{j,l} + \sum_{q=2}^Q \sum_{j=1}^t \sum_{l=-1}^2 V_{q,j,l} + W} \quad (4.6)$$

où les différents termes sont définis de la façon suivante :

- Signal utile

$$A = p_{1,1}^1 \left| \mathbf{c}_{1,1}^H \left(\mathbf{S}_{1,1}^{(0)} \right)^H \mathbf{F}_{1,1}^{(0)} \mathbf{S}_{1,1}^{(0)} \mathbf{c}_{1,1} \right|^2$$

- Interférence multi-utilisateurs

$$B = \mathbf{c}_{1,1}^H \left(\mathbf{S}_{1,1}^{(0)} \right)^H \mathbf{F}_{1,1}^{(0)} \mathbf{S}_{1,1}^{(0)} \mathbf{C}_{1,1}^2 \mathbf{P}_{1,1}^2 (\mathbf{C}_{1,1}^2)^H \left(\mathbf{S}_{1,1}^{(0)} \right)^H \left(\mathbf{F}_{1,1}^{(0)} \right)^H \mathbf{S}_{1,1}^{(0)} \mathbf{c}_{1,1}$$

- Interférence entre symboles

$$T_l = \mathbf{c}_{1,1}^H \left(\mathbf{S}_{1,1}^{(0)} \right)^H \mathbf{F}_{1,1}^{(0)} \mathbf{S}_{1,1}^{(l)} \mathbf{C}_{1,1} \mathbf{P}_{1,1} \mathbf{C}_{1,1}^H \left(\mathbf{S}_{1,1}^{(l)} \right)^H \left(\mathbf{F}_{1,1}^{(0)} \right)^H \mathbf{S}_{1,1}^{(0)} \mathbf{c}_{1,1}$$

- Interférence entre antennes d'émission

$$U_{j,l} = \mathbf{c}_{1,1}^H \left(\mathbf{S}_{1,1}^{(0)} \right)^H \mathbf{F}_{1,j}^{(l)} \mathbf{S}_{1,j}^{(l)} \mathbf{C}_{1,j} \mathbf{P}_{1,j} \mathbf{C}_{1,j}^H \left(\mathbf{S}_{1,j}^{(l)} \right)^H \left(\mathbf{F}_{1,j}^{(l)} \right)^H \mathbf{S}_{1,1}^{(0)} \mathbf{c}_{1,1}$$

- Interférence entre cellules

$$V_{q,j,l} = \mathbf{c}_{1,1}^H \left(\mathbf{S}_{1,1}^{(0)} \right)^H \mathbf{F}_{q,j}^{(l)} \mathbf{S}_{q,j}^{(l)} \mathbf{C}_{q,j} \mathbf{P}_{q,j} \mathbf{C}_{q,j}^H \left(\mathbf{S}_{q,j}^{(l)} \right)^H \left(\mathbf{F}_{q,j}^{(l)} \right)^H \mathbf{S}_{1,1}^{(0)} \mathbf{c}_{1,1}$$

- Contribution du bruit gaussien

$$W = \sigma^2 \mathbf{c}_{1,1}^H \left(\mathbf{S}_{1,1}^{(0)} \right)^H \mathcal{G} \mathcal{G}^H \mathbf{S}_{1,1}^{(0)} \mathbf{c}_{1,1}$$

4.3 Première approximation

L'expression de β_N dépend de façon assez compliquée des réalisations des différentes séquences d'étalement, il est alors assez difficile d'étudier efficacement son comportement. On se propose donc d'étudier le comportement de β_N en considérant que les nombres d'utilisateurs $K_{q,j}$ et le facteur d'étalement N tendent vers $+\infty$ en conservant un rapport de charge constant tel que $\frac{K_{q,j}}{N}$ tend vers $\alpha_{q,j}$. Cette hypothèse nous permet d'obtenir une approximation du SINR indépendante des réalisations des séquences d'étalement qui on le verra traduit assez fidèlement le comportement réel du système.

Théorème 1. *Lorsque $\forall(q, j), q = (1, \dots, Q), j = (1, \dots, t)$, $K_{q,j}$ et N tendent vers $+\infty$ dans un rapport constant, alors β_N converge en probabilité vers*

$$\beta = \frac{p_{1,1}}{\mu_{1,1}} \frac{\int_{-1/2}^{1/2} \eta(e^{2i\pi f}) df}{1 - \int_{-1/2}^{1/2} \eta(e^{2i\pi f}) df} \quad (4.7)$$

avec $\eta(e^{2i\pi f})$ défini par

$$\eta(e^{2i\pi f}) = \mu_{1,1} \mathbf{h}_{1,1}^H(e^{2i\pi f}) \left[\sum_{q=1}^Q \mathbf{H}_q(e^{2i\pi f}) \mathbf{\Gamma}_q \mathbf{H}_q^H(e^{2i\pi f}) + \sigma^2 I_r \right]^{-1} \mathbf{h}_{1,1}(e^{2i\pi f})$$

où $\mathbf{\Gamma}_q$ est une matrice diagonale de dimension $t \times t$ dont les entrées correspondent aux variances $\mu_{q,j}$ des séquences transmises par chacune des t antennes de la q^{ieme} station de base, $\mathbf{H}_q(e^{2i\pi f})$ est la valeur à la fréquence f de la transformée de Fourier du filtre $\mathbf{H}_q(z)$, et $\mathbf{h}_{1,1}(e^{2i\pi f})$ la première colonne de $\mathbf{H}_1(e^{2i\pi f})$.

La preuve du théorème précédent s'effectue en plusieurs étapes. Nous recherchons tout d'abord la limite de chacun des termes de β_N dans le régime asymptotique énoncé précédemment. Nous procéderons par la suite à la réécriture de la nouvelle expression que l'on aura obtenue.

Première étape : Étude des termes A et W

La limite des termes énoncés ci-dessus peut être obtenue en ayant recours à un lemme fréquemment utilisé dans l'analyse de performance de grands systèmes de communication (cf. [30])

Lemme 1. *Soit \mathbf{X} une matrice déterministe de dimension $N \times N$ uniformément bornée, alors*

$$\lim_{N \rightarrow +\infty} \mathbb{E} \left| \mathbf{c}_1^H \left(\mathbf{S}_{q,j}^{(0)} \right)^H \mathbf{X} \mathbf{S}_{q,j}^{(0)} \mathbf{c}_1 - \frac{1}{N} \text{Trace}(\mathbf{X}) \right|^2 = 0$$

Afin d'utiliser le lemme précédent, il est nécessaire de justifier que les matrices \mathcal{G} et $\mathbf{F}_{q,j}^{(l)}$ sont uniformément bornées, ce qui peut se faire en réutilisant les arguments utilisés dans ???. Nous sommes maintenant en mesure d'appliquer les résultats donnés par le Lemme 1, ce qui nous mène au Corollaire suivant.

Corollaire 1.

$$A - p_{1,1}^1 \left| \frac{1}{N} \text{Trace} \left(\mathbf{F}_{1,1}^{(0)} \right) \right|^2 \rightarrow 0 \quad (4.8)$$

$$W - \sigma^2 \|\mathbf{g}\|^2 \rightarrow 0 \quad (4.9)$$

où la notion de convergence correspond à celle de convergence en probabilité.

Deuxième étape : Étude des termes T_l , $U_{j,l}$ et $V_{q,j,l}$

Le comportement asymptotique de ces termes est une conséquence directe du Lemme suivant.

Lemme 2. *Soit $\Gamma = \mathbf{c}_1^H \left(\mathbf{S}_{1,1}^{(0)} \right)^H \mathbf{X} \mathbf{S}_{q,j}^{(l)} \mathbf{C}_{q,j} \mathbf{P}_{q,j} \mathbf{C}_{q,j}^H \left(\mathbf{S}_{q,j}^l \right)^H \mathbf{X}^H \mathbf{S}_{1,1}^{(0)} \mathbf{c}_1$ où \mathbf{X} est une matrice de Toeplitz déterministe de dimension $N \times N$ uniformément bornée, alors pour tout $l \neq 0$, et pour $l = 0$, lorsque $j \neq 1$ ou $q \neq 1$,*

$$\lim_{N \rightarrow +\infty} \mathbb{E} \left| \Gamma - \mu_{q,j} \frac{1}{N} \text{Trace}(\mathbf{X} \mathbf{X}^H) \right|^2 = 0$$

Nous avons déjà montré que les matrices de Toeplitz $\mathbf{F}_{q,j}^{(l)}$ sont uniformément bornées. La convergence au sens des moindres carrés impliquant celle en probabilité, nous aboutissons au Corollaire suivant.

Corollaire 2.

$$T_l - \mu_{1,1} \frac{1}{N} \text{Trace} \left(\mathbf{F}_{1,1}^{(l)} \left(\mathbf{F}_{1,1}^{(l)} \right)^H \right) \rightarrow 0 \quad (4.10)$$

$$U_{j,l} - \mu_{1,j} \frac{1}{N} \text{Trace} \left(\mathbf{F}_{1,j}^{(l)} \left(\mathbf{F}_{1,j}^{(l)} \right)^H \right) \rightarrow 0 \quad (4.11)$$

$$V_{q,j,l} - \mu_{q,j} \frac{1}{N} \text{Trace} \left(\mathbf{F}_{q,j}^{(l)} \left(\mathbf{F}_{q,j}^{(l)} \right)^H \right) \rightarrow 0 \quad (4.12)$$

où la notion de convergence correspond à celle de convergence en probabilité.

Troisième étape : Étude du terme B Le comportement asymptotique de B est une conséquence directe du Lemme suivant

Lemme 3. Soit $\Gamma = \mathbf{c}_1^H \left(\mathbf{S}_{1,1}^{(0)} \right)^H \mathbf{X} \mathbf{S}_{1,1}^{(0)} \mathbf{C}_{1,1}^2 \mathbf{P}_{1,1}^2 \left(\mathbf{C}_{1,1}^2 \right)^H \left(\mathbf{S}_{1,1}^{(0)} \right)^H \mathbf{X}^H \mathbf{S}_{1,1}^{(0)} \mathbf{c}_1$ ou \mathbf{X} est une matrice de Toeplitz déterministe de dimension $N \times N$ uniformément bornée, alors

$$\lim_{N \rightarrow +\infty} \mathbb{E} \left| \Gamma - \mu_{1,1} \left(\frac{1}{N} \text{Trace}(\mathbf{X} \mathbf{X}^H) - \left| \frac{1}{N} \text{Trace}(\mathbf{X}) \right|^2 \right) \right|^2 = 0$$

Une application directe du Lemme 3 nous conduit alors au résultat suivant.

Corollaire 3.

$$B - \mu_{1,1} \left(\frac{1}{N} \text{Trace} \left(\mathbf{F}_{1,1}^{(0)} \left(\mathbf{F}_{1,1}^{(0)} \right)^H \right) - \left| \frac{1}{N} \text{Trace} \left(\mathbf{F}_{1,1}^{(0)} \right) \right|^2 \right) \rightarrow 0 \quad (4.13)$$

où la notion de convergence correspond à celle de convergence en probabilité.

Quatrième étape : Expression du SINR

Les corollaires cités précédemment nous permettent de montrer que $\beta - \beta_N$ converge en probabilité vers 0, où

$$\beta = \frac{p_{1,1}^1}{\mu_{1,1}} \frac{\mu_{1,1} \left| \frac{1}{N} \text{Trace} \left(\mathbf{F}_{1,1}^{(0)} \right) \right|^2}{\sum_{q=1}^Q \sum_{j=1}^t \sum_{l=-1}^2 \mu_{q,j} \frac{1}{N} \text{Trace}(\mathbf{F}_{q,j}^{(l)} \left(\mathbf{F}_{q,j}^{(l)} \right)^H) + \sigma^2 \|g\|^2 - \mu_{1,1} \left| \frac{1}{N} \text{Trace} \left(\mathbf{F}_{1,1}^{(0)} \right) \right|^2}$$

Nous simplifions maintenant l'expression de β . Tout d'abord, nous remarquons que la matrice $\mathbf{F}_{1,1}^{(0)}$ est une matrice de Toeplitz de dimension N , sa trace normalisée coïncide donc avec son terme diagonal $f_1(0) = \mathbf{g}^H \mathbf{h}_{1,1}$, nous obtenons alors

$$\beta = \frac{p_{1,1}^1}{\mu_{1,1}} \frac{\mu_{1,1} |\mathbf{g}^H \mathbf{h}_{1,1}|^2}{\sum_{q=1}^Q \sum_{j=1}^t \sum_{l=-1}^2 \mu_{q,j} \frac{1}{N} \text{Trace}(\mathbf{F}_{q,j}^{(l)} \left(\mathbf{F}_{q,j}^{(l)} \right)^H) + \sigma^2 \|g\|^2 - \mu_{1,1} |\mathbf{g}^H \mathbf{h}_{1,1}|^2}$$

Ensuite, nous remarquons que

$$\sum_{l=-1}^2 \frac{1}{N} \text{Trace} \left(\mathbf{F}_{q,j}^{(l)} \left(\mathbf{F}_{q,j}^{(l)} \right)^H \right) = \frac{1}{N} \text{Trace} (\mathcal{F}_{q,j} \mathcal{F}_{q,j}^H)$$

Comme $\mathcal{F}_{q,j} \mathcal{F}_{q,j}^H$ est une matrice de Toeplitz de dimension N , sa trace normalisée correspond à son terme diagonale $\|f_{q,j}(0)\|^2$. Sachant que $f_{q,j}(0) = \mathbf{g}^H \mathcal{H}_{q,j}$, nous obtenons

$$\frac{1}{N} \text{Trace} (\mathcal{F}_{q,j} \mathcal{F}_{q,j}^H) = \mathbf{g}^H \mathcal{H}_{q,j} \mathcal{H}_{q,j}^H \mathbf{g}$$

Ainsi nous réécrivons β de la façon suivante

$$\beta = \frac{p_{1,1}^1}{\mu_{1,1}} \frac{\mu_{1,1} |\mathbf{g}^H \mathbf{h}_{1,1}|^2}{\sum_{q=1}^Q \sum_{j=1}^t \mu_{q,j} \mathbf{g}^H \mathcal{H}_{q,j} \mathcal{H}_{q,j}^H \mathbf{g} + \sigma^2 \|g\|^2 - \mu_{1,1} |\mathbf{g}^H \mathbf{h}_{1,1}|^2}$$

En utilisant l'expression de \mathbf{g} , donnée par (3.6), nous obtenons que

$$\sum_{q=1}^Q \sum_{j=1}^t \mu_{q,j} \mathbf{g}^H \mathcal{H}_{q,j} \mathcal{H}_{q,j}^H \mathbf{g} + \sigma^2 \|g\|^2 = \mathbf{g}^H \mathbf{h}_{1,1}$$

Ces considérations nous mènent à une expression plus simple de β , soit

$$\beta = \frac{p_{1,1}^1}{\mu_{1,1}} \frac{\mu_{1,1} \mathbf{g}^H \mathbf{h}_{1,1}}{1 - \mu_{1,1} \mathbf{g}^H \mathbf{h}_{1,1}}$$

Nous posons alors

$$\eta = \mu_{1,1} \mathbf{g}^H \mathbf{h}_{1,1} \quad (4.14)$$

$$= \mu_{1,1} \mathbf{h}_{1,1}^H \left(\sum_{q=1}^Q \sum_{j=1}^t \mu_{q,j} \mathcal{H}_{q,j} \mathcal{H}_{q,j}^H + \sigma^2 \mathbf{I}_r \right)^{-1} \mathbf{h}_{1,1} \quad (4.15)$$

On achève la démonstration en remarquant que la densité spectrale de puissance associée à la matrice de Toeplitz $\mu_{q,j} \mathcal{H}_{q,j} \mathcal{H}_{q,j}^H$ est égale à $\mu_{q,j} \mathbf{h}_{q,j}(e^{2i\pi f}) \mathbf{h}_{q,j}^H(e^{2i\pi f})$, ce qui entraîne en grande dimension l'approximation suivante

$$\eta \approx \int_{-1/2}^{1/2} \eta(e^{2i\pi f}) df$$

où

$$\eta(e^{2i\pi f}) = \mu_{1,1} \mathbf{h}_{1,1}^H(e^{2i\pi f}) \left[\sum_{q=1}^Q \mathbf{H}_q(e^{2i\pi f}) \mathbf{\Gamma}_q \mathbf{H}_q^H(e^{2i\pi f}) + \sigma^2 \mathbf{I}_r \right]^{-1} \mathbf{h}_{1,1}(e^{2i\pi f})$$

4.4 Seconde approximation

Dans la section suivante, on se propose de déterminer une approximation de β à l'aide de termes déterministes indépendants des réalisations particulières du canal de propagation. Pour ce faire, on se place dans le cas où les nombres d'antennes en émission et en réception tendent vers l'infini en conservant un rapport $\frac{t}{r}$ constant.

L'allure de l'expression de β et l'observation que la matrice de covariance de $\text{vec}\left(\mathbf{H}_q(e^{2i\pi f})\mathbf{\Gamma}_q^{1/2}\right)$ correspond à $\frac{a_q}{t}\mathbf{\Gamma}_q \otimes I_r$ nous pousse à introduire un canal MIMO équivalent $\mathcal{H}(e^{2i\pi f})$ de dimension $r \times Qt$ donné par

$$\mathcal{H}(e^{2i\pi f}) = \tilde{\mathcal{H}}(e^{2i\pi f})\tilde{\mathcal{C}}^{1/2} \quad (4.16)$$

où l'on définit par

– $\tilde{\mathcal{C}}$, une matrice bloc diagonale de dimension $Qt \times Qt$ donnée par

$$\tilde{\mathcal{C}} = \begin{bmatrix} a_1\mathbf{\Gamma}_1 & & 0 \\ 0 & \ddots & 0 \\ 0 & & a_Q\mathbf{\Gamma}_Q \end{bmatrix}$$

Nous rappelons que les $a_q, q = 1, \dots, Q$ correspondent aux coefficients de pondération de la puissance reçue en provenance des différentes stations de bases par rapport à la puissance de référence de la station principale. Les matrices $\mathbf{\Gamma}_q$ de dimension $t \times t$ sont diagonales et leurs entrées correspondent aux variances des flux transmis par les différentes antennes d'émission de la q^{ieme} station de base, elles sont données par

$$\mathbf{\Gamma}_q = \begin{bmatrix} \mu_{q,1} & & 0 \\ 0 & \ddots & 0 \\ 0 & & \mu_{q,t} \end{bmatrix}$$

– $\tilde{\mathcal{H}}(e^{2i\pi f})$, une matrice aléatoire gaussienne de dimension $r \times Qt$ dont les entrées sont des variables i.i.d. centrées de variance $\frac{1}{t}$

Dans ces conditions, nous pouvons énoncer le résultat suivant

Théorème 2. *Lorsque t et r tendent vers $+\infty$ dans un rapport constant, β converge en probabilité vers*

$$\beta^* = \frac{p_{1,1}^1}{\mu_{1,1}} \left[\frac{1}{\left(I_{Qt} + \delta\tilde{\mathcal{C}}\right)_{1,1}} - 1 \right] = p_{1,1}^1\delta \quad (4.17)$$

où δ est la solution unique strictement positive de l'équation suivante

$$\delta = \frac{r}{t} \frac{1}{\sigma^2 + \frac{1}{t} \text{Trace} \left[\tilde{\mathcal{C}} \left(I_{Qt} + \delta\tilde{\mathcal{C}}\right)^{-1} \right]} \quad (4.18)$$

Afin de justifier le théorème précédent, il est nécessaire de réécrire différemment l'expression de $\eta(e^{2i\pi f})$. L'utilisation du lemme d'inversion matricielle et de l'expression des entrées diagonales de l'inverse d'une matrice positive nous permet de réécrire $\eta(e^{2i\pi f})$ de la façon suivante

$$\eta(e^{2i\pi f}) = 1 - \left(\tilde{\mathbf{Q}}(e^{2i\pi f}) \right)_{1,1}$$

où $\tilde{\mathbf{Q}}(e^{2i\pi f})$ est définie par

$$\tilde{\mathbf{Q}}(e^{2i\pi f}) = \sigma^2 \left(\tilde{\mathcal{C}}^{\frac{1}{2}} \tilde{\mathcal{H}}(e^{2i\pi f})^H \tilde{\mathcal{H}}(e^{2i\pi f}) \tilde{\mathcal{C}}^{\frac{1}{2}} + \sigma^2 \mathbf{I}_{Q_t} \right)^{-1}$$

Par ailleurs, pour chaque valeur de f , la matrice $\tilde{\mathcal{H}}(e^{2i\pi f})$ de dimension $r \times Q_t$ est à entrées i.i.d. centrées de variance $\frac{1}{t}$. Dans ces conditions, nous pouvons déduire du chapitre 8 que $\left(\tilde{\mathbf{Q}}(e^{2i\pi f}) \right)_{1,1}$ converge surement vers $\left(I_{Q_t} + \delta \tilde{\mathcal{C}} \right)_{1,1}^{-1}$ où δ est l'unique solution strictement positive de l'équation

$$\delta = \frac{r}{t} \frac{1}{\sigma^2 + \frac{1}{t} \text{Trace} \left[\tilde{\mathcal{C}} \left(I_{Q_t} + \delta \tilde{\mathcal{C}} \right)^{-1} \right]} \quad (4.19)$$

Dans ces conditions,

$$\frac{\int_{-1/2}^{1/2} \eta(e^{2i\pi f}) df}{1 - \int_{-1/2}^{1/2} \eta(e^{2i\pi f}) df}$$

converge vers

$$\frac{1 - \left(I_{Q_t} + \delta \tilde{\mathcal{C}} \right)_{1,1}^{-1}}{\left(I_{Q_t} + \delta \tilde{\mathcal{C}} \right)_{1,1}^{-1}} = \frac{1}{\left(I_{Q_t} + \delta \tilde{\mathcal{C}} \right)_{1,1}^{-1}} - 1 = a_1 \mu_{1,1} \delta \quad (4.20)$$

Par conséquent et en rappelant que $a_1 = 1$, β converge vers $\beta^* = p_{1,1}^1 \delta$.

4.5 Simulations

Dans cette section nous présentons les résultats de simulations des quatre scénarios envisagés précédemment dans le but d'illustrer la pertinence de β et β^* , les approximations asymptotiques du SINR réel β_N , dont les formules sont respectivement données par (4.7), (4.17) et (4.6).

Les paramètres communs à l'ensemble des simulations sont présentés maintenant. Nous considérons que le rayon des cellules est égal à $R = 0.577$ kilomètres. Les stations de base sont équipées de $t = 4$ antennes de transmissions et les mobiles de $r = 4$ antennes de réception. Les canaux de propagation situés entre chaque station de base et le mobile, sujets aux phénomènes d'évanouissement et de sélectivité fréquentielle, ont été générés selon le modèle gaussien présenté précédemment. La longueur de la réponse impulsionnelle commune à chacun d'eux a été fixée à $L = 5$. Finalement nous considérons une gamme de $\frac{E_b}{N_0}$ allant de -15 dB à 10 dB.

4.5.1 Présentation des scénarios de simulation

Scénario 1 : Aucune interférences entre cellules

Pour ce scénario, comme l'illustre la figure 4.1, nous fixons le nombre de stations de base Q à 1. Implicitement, nous ne considérons donc pas d'interférences entre cellules. Le mobile test se trouve alors à une distance $d = 0.385$ kilomètres de la station de base.

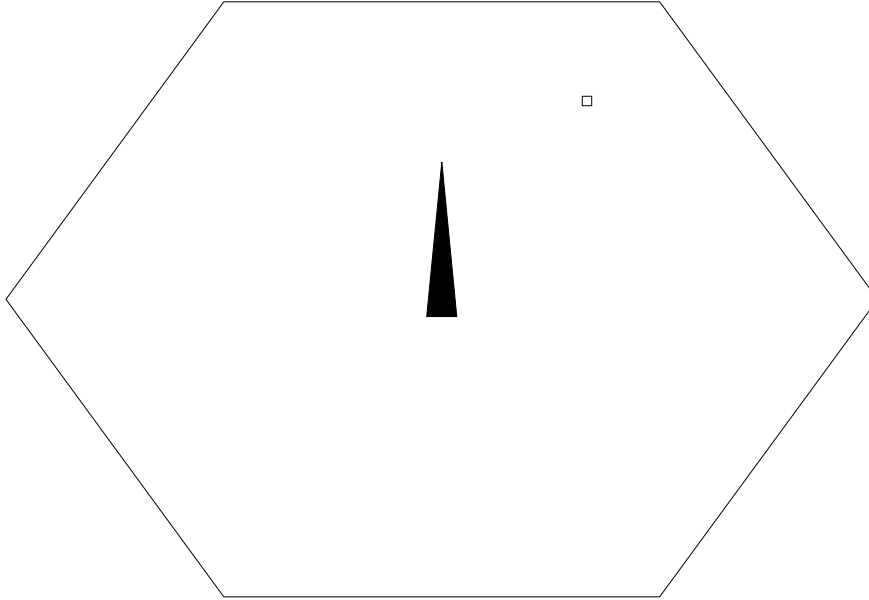


FIGURE 4.1 – Scénario 1, géométrie du réseau

Scénario 2 : A l'intérieur de la cellule

Pour ce scénario, le nombre de stations de base est fixé à $Q = 8$. La figure 4.2 représente l'agencement des cellules et le tableau 4.1 rassemble les distances en kilomètres entre le mobile et les différentes stations de bases. Tout comme lors du scénario 1, le mobile se situe à l'intérieur de sa cellule à une distance de $d = 0.385$ kilomètres de la station de base principale.

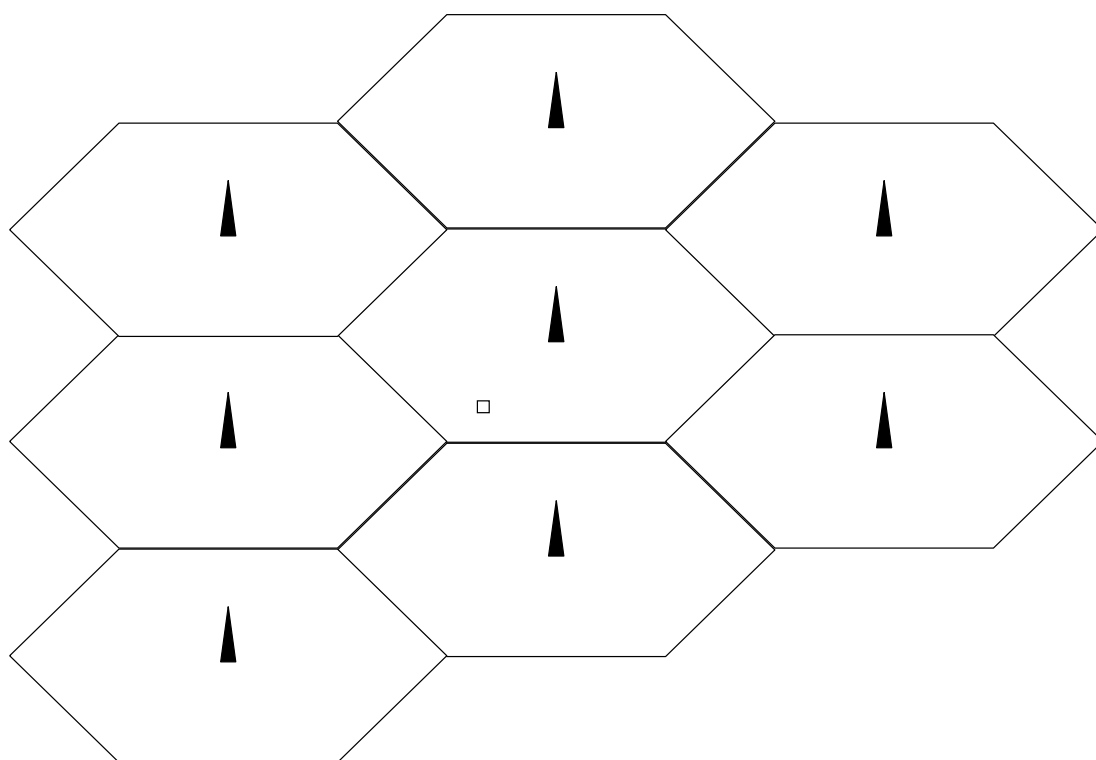


FIGURE 4.2 – Scénario 2, géométrie du réseau

BTS n^o	Distance Mobile-BTS (km)
1	0.385
2	2.08
3	2.08
4	3.212
5	3.212
6	4.039
7	4.039
8	4.039

TABLE 4.1 – Scénario 2 : Distances Mobile - Stations de Base

Scénario 3 : A l'intersection de 2 cellules

Pour le scénario 3, le nombre de stations de base est également fixé à $Q = 8$. La figure 4.3 représente l'agencement des cellules et le tableau 4.2 rassemble les distances en kilomètres entre le mobile et les différentes stations de bases. Cette fois-ci, le mobile se situe à l'intersection de deux cellules à une distance de $d = 0.5$ kilomètres de la station de base principale.

BTS n^o	Distance Mobile-BTS (km)
1	0.5
2	0.5
3	0.865
4	0.865
5	1.322
6	1.322
7	1.322
8	1.322

TABLE 4.2 – Scénario 3 : Distances Mobile - Stations de Base

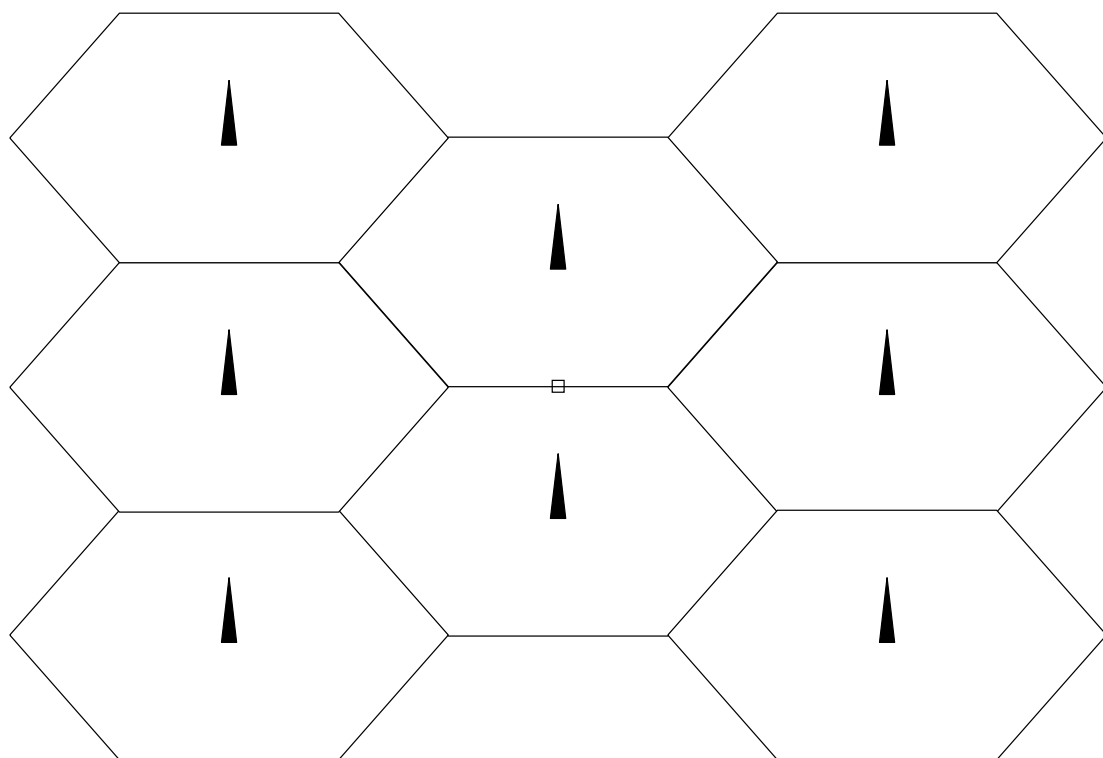


FIGURE 4.3 – Scénario 3, géométrie du réseau

Scénario 4 : A l'intersection de 3 cellules

Pour le scénario 4, le nombre de stations de base est fixé à $Q = 9$. La figure 4.4 représente l'agencement des cellules et le tableau 4.3 rassemble les distances en kilomètres entre le mobile et les différentes stations de bases. Cette fois-ci, le mobile se situe à l'intersection de trois cellules à une distance de $d = 0.577$ kilomètres de la station de base principale.

BTS n^o	Distance Mobile-BTS (km)
1	0.577
2	0.577
3	0.577
4	1.154
5	1.154
6	1.154
7	1.526
8	1.526
9	1.526

TABLE 4.3 – Scénario 4 : Distances Mobile - Stations de Base

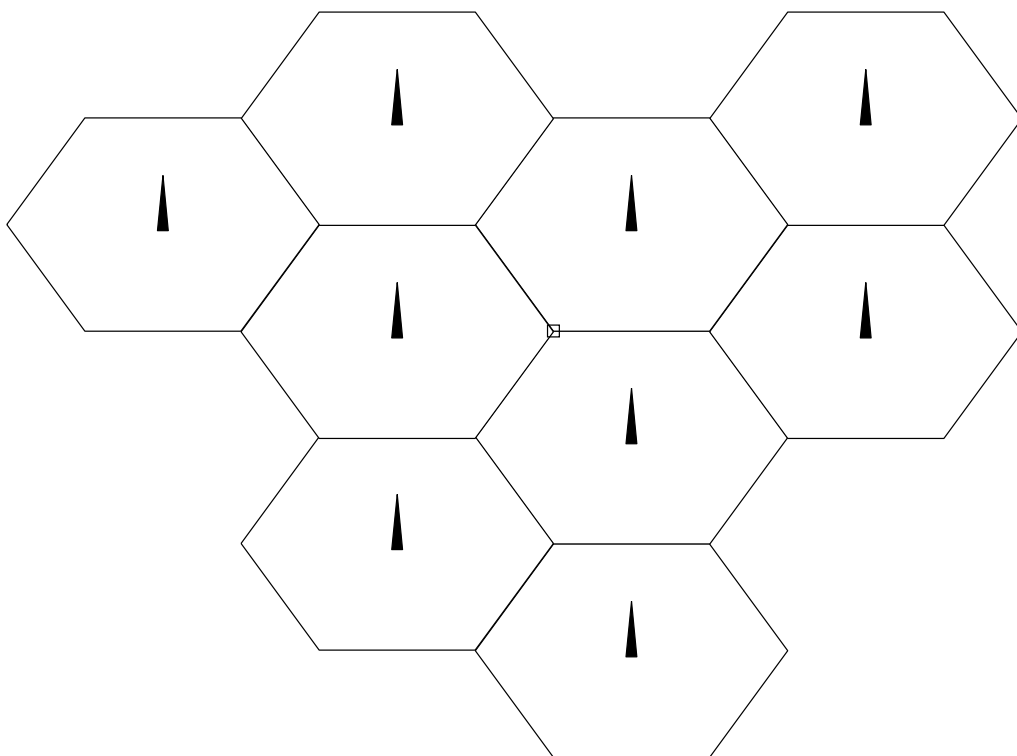


FIGURE 4.4 – Scénario 4, géométrie du réseau

4.5.2 Validation des approximations

Allocation de puissance uniforme

Dans un premier temps, nous nous intéressons au scénario 1 pour lequel aucune source d'interférence extra-cellulaire ne vient perturber le signal reçu par le mobile, alors seulement affecté par le bruit thermique inhérent au matériel de réception. Pour ce scénario, le nombre de stations de base est donc fixé à $Q = 1$. Nous plaçons le mobile à l'intérieur de la cellule à une distance $d = 0.385$ kilomètre et nous considérons que l'allocation de puissance est uniforme. En d'autres termes, les puissances $p_{1,j}^i, i = 1, \dots, K_{1,j}, j = 1, \dots, t$ allouées à chacun des utilisateurs de chacune des antennes d'émission de l'unique station de base sont toutes égales.

La figure 4.5 représente le taux d'erreur binaire en sortie du récepteur de Wiener dans le cadre du scénario 1 lorsque l'allocation de puissance est uniforme. Les performances réelles obtenues à partir de l'expression de β_N , donné par l'équation 4.6, ont été calculées à l'aide d'une boucle de Monte Carlo sur 100 réalisations du canal de propagation pour différentes valeurs du nombre d'utilisateurs $K_{1,1}$ (16 et 8) et du facteur d'étalement N (64,32) en veillant à conserver un rapport de charge constant fixé à $\alpha_{1,1} = \frac{K_{1,1}}{N} = \frac{1}{4}$. Pour chacune des réalisations du canal de propagation nous avons alors calculé le taux d'erreur binaire moyen de chacun des symboles d'un bloc en contenant 30 au total en faisant varier pour chacun d'eux les réalisations du code de scrambling. Les performances théoriques données par β , la première approximation de β_N , ont été calculé à partir de l'équation 4.7 également par simulation de Monte-Carlo avec les mêmes réalisations du canal de propagation que pour le calcul des performances réelles. Enfin, les performances théoriques données par β^* , la seconde approximation de β_N ont été calculées à l'aide de l'équation 4.17.

Tout d'abord, nous constatons que les courbes de performances réelles sont pratiquement confondues. Nous en déduisons donc que les performances du récepteur de Wiener sous-optimal ne dépendent pas du nombre d'utilisateurs $K_{1,1}$ ou du facteur d'étalement N , mais seulement du rapport de charge $\alpha_{1,1}$. Ensuite nous pouvons remarquer que la courbe de performance issue de la première approximation du SINR est pratiquement confondue avec les courbes de performances réelles. β constitue donc une excellente approximation de β_N dans le cadre d'une allocation de puissance uniforme et dans le contexte d'un environnement libre de toute interférences inter-cellulaires. Nous pouvons tout de même constater un écart entre les perfor-

mances correspondant aux SINR β_N et β_\star qui croit raisonnablement à mesure que le rapport $\frac{E_b}{N_0}$ augmente. Ceci s'explique simplement par le faible nombre d'antennes considérées en émission et en réception.

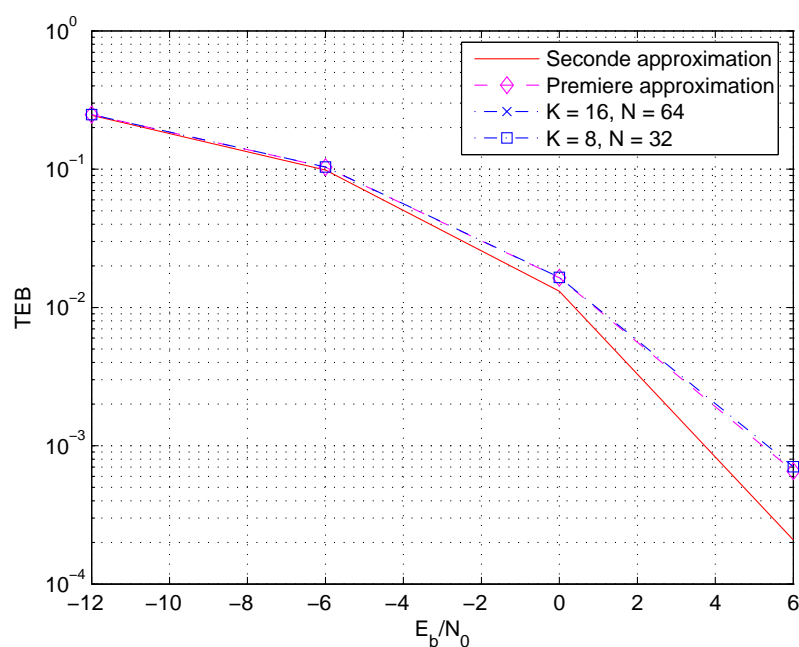


FIGURE 4.5 – Scénario 1, puissance uniforme

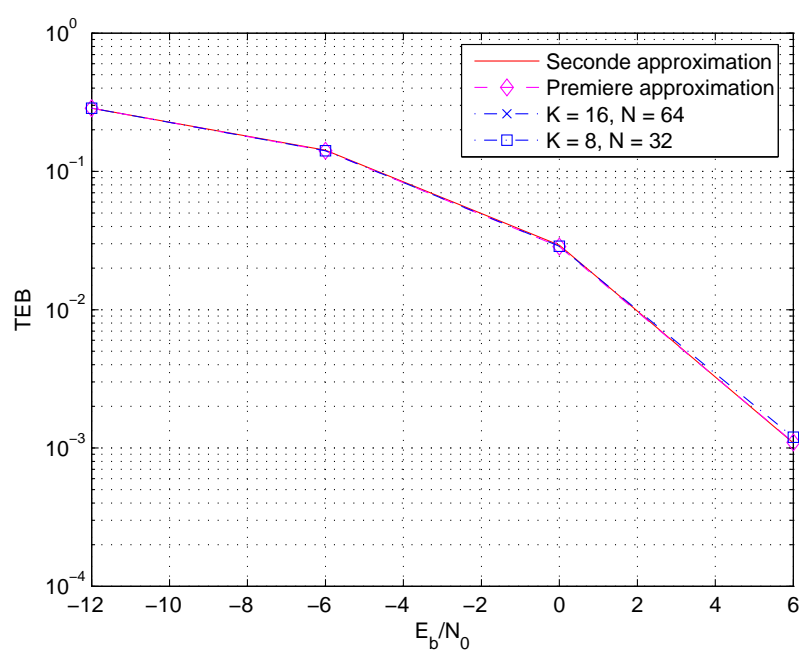


FIGURE 4.6 – Scénario 2, puissance uniforme

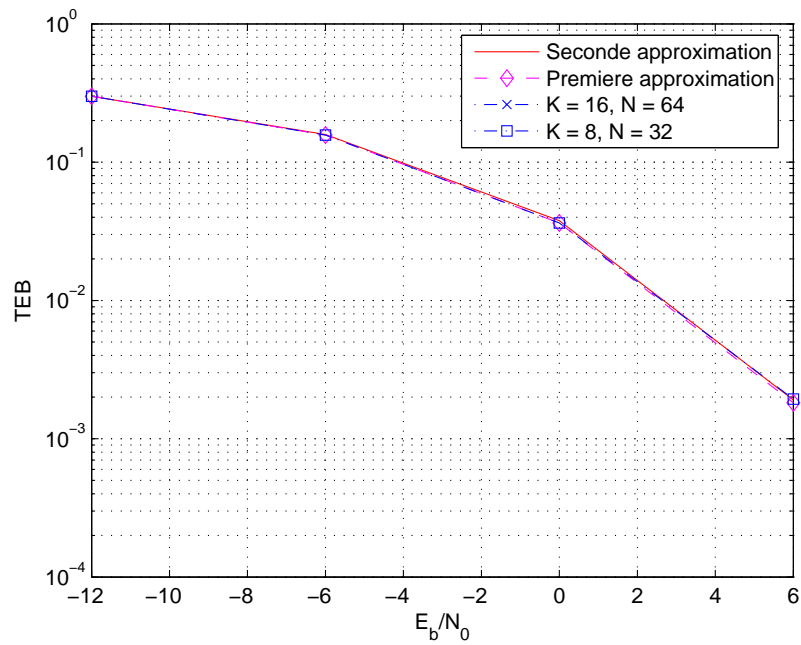


FIGURE 4.7 – Scénario 3, puissance uniforme

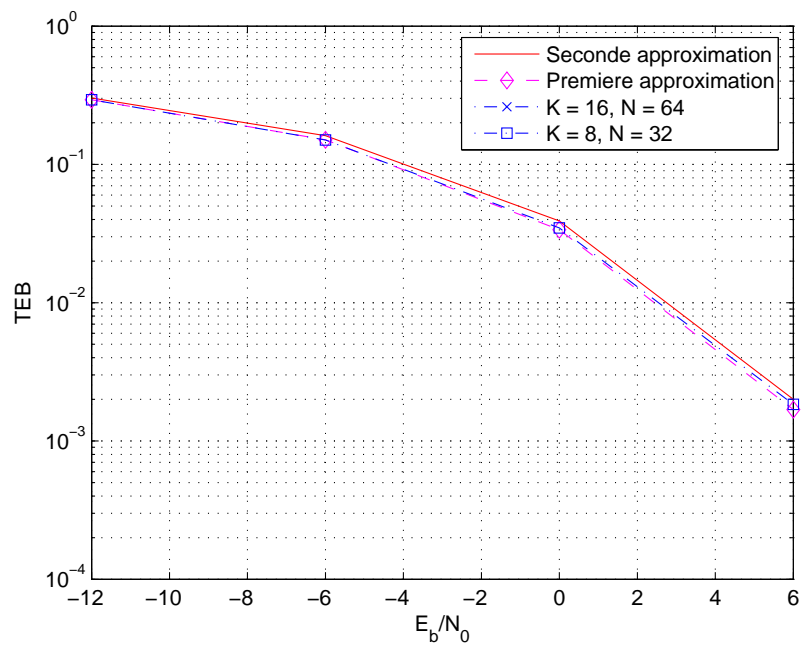


FIGURE 4.8 – Scénario 4, puissance uniforme

Dans le cadre du scénario 1, nous avons vu que l'erreur commise en approximant le SINR β_N par β^* peut être assez importante. Heureusement, l'intégration des interférences inter-cellulaires à notre modèle va avoir pour impact de limiter l'erreur commise, ce que nous allons voir tout de suite à travers les résultats de simulations des scénarios 2,3 et 4. Rappelons que pour ces scénarios, nous considérons le cas pour lequel le signal reçu au niveau du mobile est de plus en plus perturbé par les cellules adjacentes à la cellule principale. Les figures 4.6, 4.7 et 4.8 représentent respectivement le taux d'erreur binaire en sortie du récepteur de Wiener pour les scénarios 2,3 et 4 dans les mêmes conditions que celles qui nous ont menées à la figure 4.5. Pour l'ensemble de ces simulations, le niveau de puissance des cellules interférentes a été fixé au maximum délivrable, soit 43 dBm.

Nous constatons encore une fois, et pour l'ensemble des scénarios, que les différentes courbes de performances réelles sont confondues. Ceci nous confirme alors que, à puissance comparable, les performances du filtre de Wiener dépendent uniquement de la charge $\alpha_{1,1}$. Nous remarquons également que les courbes de performances données par la première approximation du SINR sont extrêmement proches des courbes de performances réelles, ce qui nous assure encore ici de la pertinence de β . Ensuite, contrairement aux résultats du scénario 1, nous constatons que l'écart entre les courbes réelles et celles obtenues à l'aide de β^* est vraiment faible même lorsque le rapport $\frac{E_b}{N_0}$ est élevé. Il est clair que β^* est une bien meilleure approximation de β_N lorsque l'on inclut les effets d'interférences entre cellules. En effet, l'augmentation du nombre d'antennes émettrices "accélère" la convergence du taux d'erreur moyen vers sa moyenne statistique.

Allocation de puissance non uniforme

Afin d'évaluer la pertinence de nos approximations dans le cas où l'énergie bit allouée à chacun des utilisateurs est différentes, nous avons réparti les $K_{1,1}$ utilisateurs parmi trois classes d'allocation de puissance correspondant respectivement aux niveaux $-3dB, 0dB$ et $+3dB$. Plus précisément, nous avons affecté 6 utilisateurs à la première classe, 7 à la seconde et 3 à la troisième. La distance mobile à station de base étant la principale cause d'atténuation de puissance, il est légitime de considérer que les utilisateurs les plus lointains appartiennent à la troisième classe alors que les plus proches appartiennent à la première classe, en d'autres termes plus l'utilisateur est loin plus il se voit attribuer un niveau de puissance élevé. Le

facteur d'étalement N commun à l'ensemble des utilisateurs a été fixé à 64. Nous considérons à nouveau le scénario 1 pour lequel aucune interférence extra cellulaire ne vient perturber le signal.

La figure 4.9 représente les taux d'erreur binaire (TEB) obtenus avec nos différentes expressions du SINR en fonction du rapport $\frac{E_b}{N_0}$ pour chacune des trois classes de puissance. Tout d'abord, nous pouvons constater qu'à rapport signal à bruit équivalent, les performances réelles et théoriques de chacune des classes sont différentes. Ceci illustre la sensibilité du récepteur de Wiener sous-optimal aux interférences multi-utilisateurs. De plus nous constatons que les performances des utilisateurs disposant de moins de puissance sont plus faibles. Ceci illustre un phénomène propre au CDMA appelé communément effet de masque, que l'on explique de la façon suivante : les utilisateurs les plus lointains sont peu affectés par la puissance allouée aux utilisateurs les plus proche ; A contrario, la puissance allouée aux utilisateurs lointains constitue une forte source d'interférence pour les utilisateurs les plus proches. En ce qui concerne l'écart entre performances théoriques et réelles, nous pouvons constater que la première approximation du SINR β traduit encore une fois parfaitement le SINR réel β_N et ce pour chacune des 3 classes de puissances. En revanche, comme nous avons pu l'observer dans le cadre de l'allocation de puissance uniforme, l'écart entre les performances réelles et celles données par la seconde approximation β^* grandit à mesure que le SNR augmente.

Nous considérons maintenant le même type d'allocation de puissance par classe dans le cadre des scénarios 2,3 et 4. Les figures 4.10, 4.11 et 4.12 présentent alors les performances théoriques et réelles de chacun de ces scénarios pour chacune des classe de puissance. Nous pouvons à nouveau constater la relative faiblesse du filtre de Wiener sous-optimal aux interférences multi-utilisateurs. En effet, la dispersion des performances pour chacune des classes est ici encore plus importante que pour le scénario 1, et d'autant plus importante que les sources d'interférences sont puissantes. Également, pour chacun de ces scénarios nous pouvons constater l'impact de l'effet de masque sur les performances des utilisateurs appartenant à la première classe et à la seconde classe.

En ce qui concerne l'écart entre prévisions théoriques et performances réelles, tout d'abord nous remarquons encore une fois l'excellente fiabilité de β dont la courbe de performance associée se confond avec celles des performances réelles données par β_N . Ensuite, comme nous l'avions remarqué dans le cadre d'une allocation de puissance

uniforme, l'ajout de sources d'interférences extra cellulaire améliore très nettement la fiabilité de la seconde approximation. En effet, pour les scénarii 2,3 et 4 les courbes de performances données par β_N et celles données par β^* du SINR sont confondues.

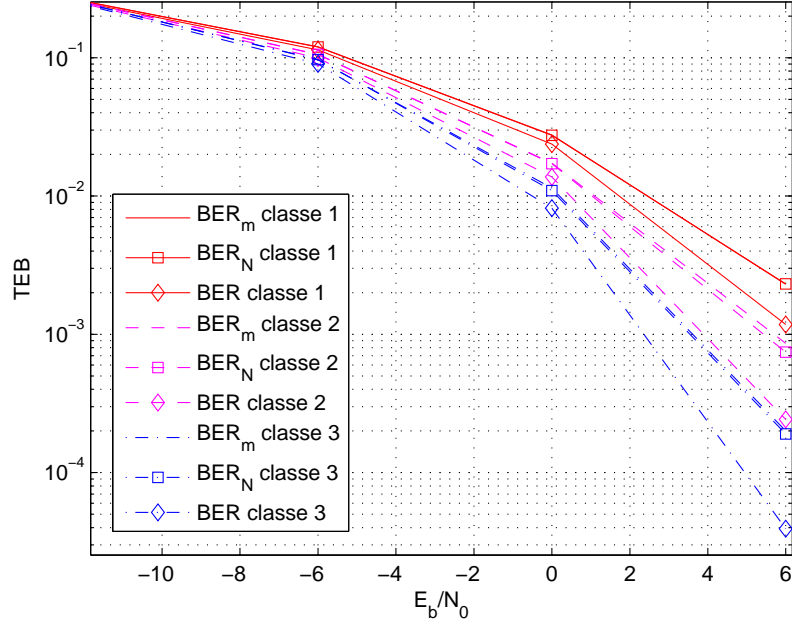


FIGURE 4.9 – Scénario 1, Distribution des puissances non uniforme

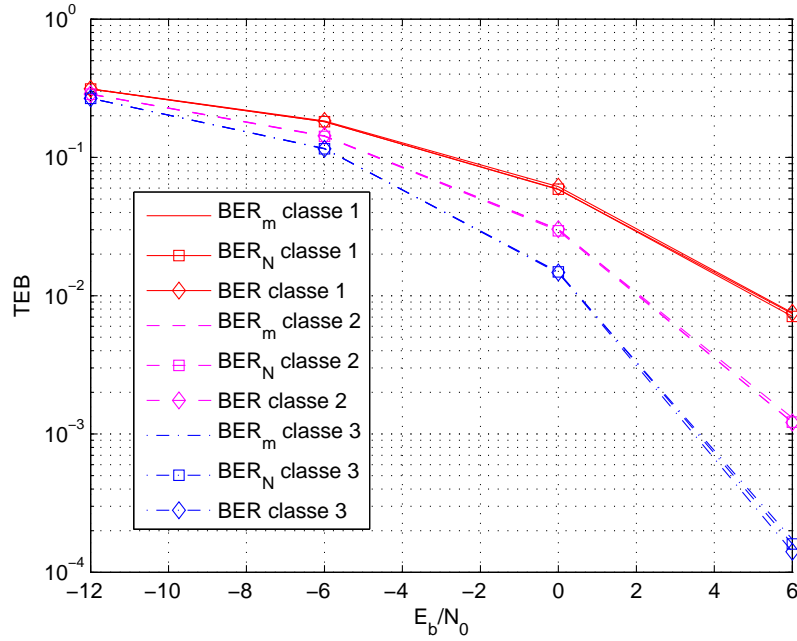


FIGURE 4.10 – Scénario 2, Distribution des puissances non uniforme

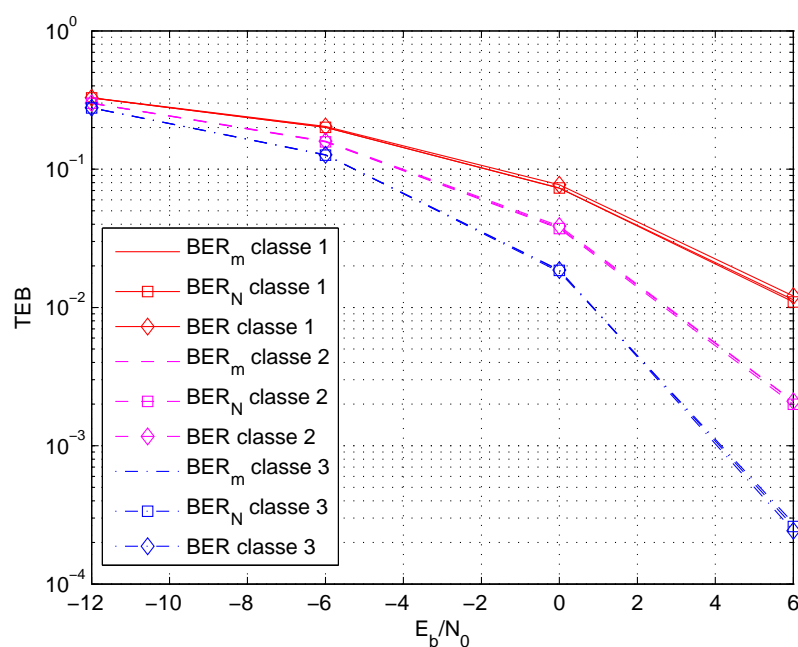


FIGURE 4.11 – Scénario 3, Distribution des puissances non uniforme

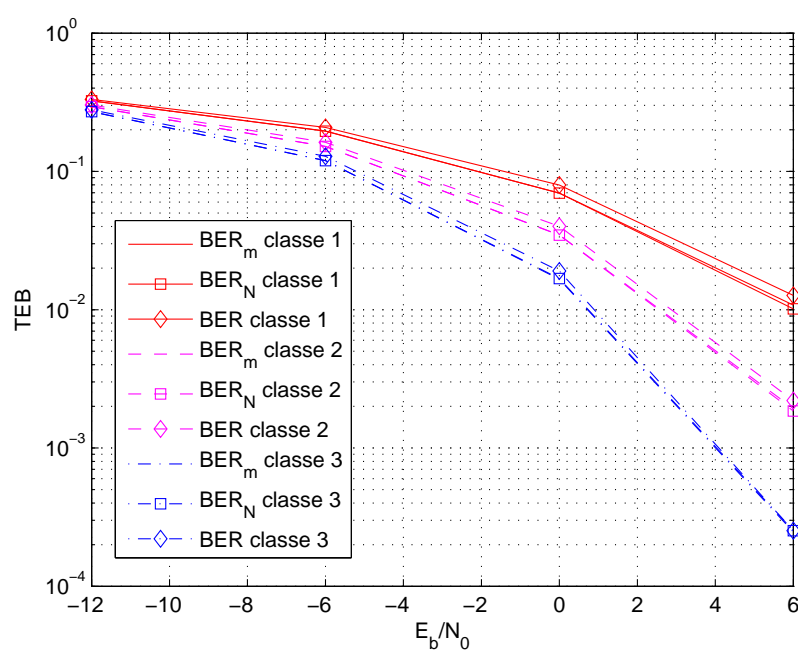


FIGURE 4.12 – Scénario 4, Distribution des puissances non uniforme

Chapitre 5

Contrôle de puissance à long terme

Dans cette partie nous utilisons les expressions asymptotiques du SINR obtenues précédemment afin de résoudre un problème d'allocation de puissance ayant pour but de maintenir une certaine qualité de lien entre une station de base et ses utilisateurs.

Dans un premier temps nous présentons l'idée générale relative à notre méthode d'allocation de puissance. Ensuite nous présenterons succinctement le modèle des signaux que nous considérons puis nous appliquerons les résultats obtenus dans le chapitre précédent en ce qui concerne les performances asymptotiques du récepteur de Wiener sous-optimal. Enfin nous décrirons notre algorithme d'allocation de puissance à long terme et donnerons des résultats de simulations permettant de s'assurer de la validité de la méthode.

5.1 Idée générale

Notre problème d'allocation consiste à rechercher les valeurs des puissances à allouer aux différents utilisateurs de la cellule au niveau de la station de base afin que le taux d'erreur binaire moyen en sortie du récepteur de Wiener sous-optimal qui leur est associé soit égal à une certaine valeur cible. Afin de relier ce travail à ce que l'on peut trouver dans un contexte réel, nous considérons que la puissance totale pouvant être transmise par la station de base est limitée à un certain nombre de Watts (20W étant une valeur classique). A taille de cellule fixée, cette hypothèse nous impose de choisir intelligemment la valeur de BER cible afin que celle-ci soit atteignable par l'ensemble des utilisateurs et ce pour la majorité des distributions spatiales de ceux-ci au sein de la cellule. Lors de la mise en oeuvre de notre algorithme, nous choisirons

donc le BER cible à l'aide d'un bilan de liaison inspiré de [25]. Afin d'estimer les taux d'erreurs binaires nous avons à nouveau recours à la fonction d'erreur $Q(\sqrt{SINR})$. Nous appelons alors β_{cible} le SINR en sortie du récepteur de Wiener de chacun des utilisateurs leur permettant d'accéder au taux d'erreur moyen voulu. En considérant un nombre d'utilisateur égal à K et en notant $\beta_i(p_1, \dots, p_K)$ le SINR réel en sortie du récepteur de Wiener associé au i^{me} utilisateur que l'on obtient avec une allocation de puissance (p_1, \dots, p_K) , le problème ci-dessus revient à satisfaire pour toute valeur de i l'égalité suivante

$$\mathbb{E} \left[Q \left(\sqrt{\beta_i(p_1, \dots, p_K)} \right) \right] = Q \left(\sqrt{\beta_{cible}} \right) \quad (5.1)$$

Il est à noter que ici l'espérance mathématique porte sur les réalisations des codes de scrambling de chacune des antennes du réseau d'émission ainsi que les différentes réalisations du canal de propagation MIMO.

Déterminer en un temps raisonnable le jeu de puissance assurant l'égalité exprimée ci-dessus n'est pas chose aisée. Solutionner de façon directe ce problème d'optimisation nécessite de faire intervenir une boucle de Monte Carlo sur les différentes réalisations des codes de scrambling et des canaux de propagation afin de calculer l'espérance mathématique du taux d'erreur binaire. Entre autres, il faudrait calculer le SINR réel à chaque itération de la boucle de Monte-Carlo pour en déterminer le taux d'erreur binaire correspondant à l'aide de la fonction d'erreur binaire puis en calculer la moyenne sur l'ensemble des itérations. Cette trop grande complexité de calcul nous dissuade de procéder de la sorte.

Ici intervient l'approximation déterministe du SINR que nous avons développé dans le chapitre précédent. On se propose d'approximer le taux d'erreur binaire moyen en sortie du récepteur de Wiener sous-optimal par le taux d'erreur binaire que l'on peut espérer avoir lorsque le nombre d'utilisateurs et le facteur d'étalement tendent vers l'infini dans un rapport constant et que les nombres d'antennes en émission et en réception tendent vers l'infini de la même façon. En effet nous avons pu voir que le comportement du SINR en sortie du récepteur de Wiener dans de telles conditions pouvait être approximé à l'aide d'une fonction déterministe. En appelant $\beta_i^*(p_1, \dots, p_K)$ l'approximation du SINR en sortie du filtre de Wiener sous-optimal associé au i^{ieme} utilisateur correspondant à ce taux d'erreur d'erreur binaire, nous pouvons alors écrire l'approximation suivante

$$\mathbb{E} \left[Q \left(\sqrt{\beta_i(p_1, \dots, p_K)} \right) \right] \approx Q \left(\sqrt{\beta_i^*(p_1, \dots, p_K)} \right)$$

ce qui nous permet de simplifier le problème d'allocation en réécrivant l'équation 5.1 de la façon suivante

$$Q(\sqrt{\beta_i^*(p_1, \dots, p_K)}) = Q(\sqrt{\beta_{cible}})$$

La nature déterministe de β_i^* va ainsi nous permettre de nous affranchir de la boucle de Monte Carlo afin de résoudre notre problème d'allocation de puissance, réduisant de ce fait énormément la complexité de calcul nécessaire. Afin de déterminer la solution de ce problème nous allons tout d'abord préciser le modèle des signaux et donner l'expression de $\beta_i^*(p_1, \dots, p_K), i = 1, \dots, K$.

5.2 Modèle des signaux

Nous considérons le cas pour lequel aucune interférence extra-cellulaire ne vient perturber le signal reçu au niveau du réseau d'antennes équipant chacun des mobiles, par conséquent le nombre de stations de bases Q est fixé à 1. Chacune des antennes d'émission de la station de base est alors chargée de transmettre K suites de symboles indépendantes en direction des K utilisateurs que contient la cellule.

L'expression du signal reçu à l'instant chip n au niveau du réseau d'antenne de réception du i^{eme} utilisateur est alors donnée par

$$\mathbf{y}_i(n) = \sum_{l=0}^{L-1} \mathbf{H}_i^l \mathbf{d}(n-l) + \mathbf{v}_i(n)$$

où l'on définit par

- \mathbf{H}_i^l , une matrice de dimension $r \times t$ dont les entrées représentent la l^{ieme} valeur des gains complexes du canal de propagation MIMO située entre le réseau d'antenne de la station de base et celui du i^{ieme} utilisateur
- $\mathbf{d}(n)$, un vecteur de dimension t correspondant à la valeur prise par le signal transmis à l'instant chip n par la j^{ieme} antenne d'émission.
- $\mathbf{v}_i(n)$, un vecteur de dimension r correspondant au bruit thermique associé au récepteur du i^{ieme} utilisateur. Ce bruit est considéré comme étant blanc et gaussien de variance σ_i^2

Nous rappelons que les séquences $\mathbf{d}(n)$, construites sur le modèle du CDMA, correspondent la somme de K suites indépendantes de symboles issus d'une constellation QAM4. La concaténation de N éléments consécutifs de la séquence $\mathbf{d}(n)$ nous

permet alors d'exprimer sous forme matricielle les opérations d'allocation de puissance, d'étalement et de scrambling opérées sur les K symboles transmis à l'instant symbole m . En appelant $\mathbf{D}_j(m)$ cette concaténation, nous pouvons alors écrire

$$\mathbf{D}_j(m) = \mathbf{S}_j^{(m)} \mathbf{C}_j \mathbf{P} \mathbf{B}_j(m)$$

où l'on définit par

- $\mathbf{S}_j^{(l)}$, une matrice diagonale de dimension $N \times N$ dont les entrées correspondent à la portion de la séquence d'embrouillage associée à l'antenne d'émission j à l'instant symbole m .
- \mathbf{C}_j , une matrice d'Hadamard de dimension $N \times K$ dont les colonnes correspondent au code d'étalement associé à chacun des K utilisateurs
- \mathbf{P} , une matrice diagonale de dimension $K \times K$ dont les entrées correspondent à la puissance allouée à chacune des K suite de symboles
- $\mathbf{B}_j(m)$, un vecteur de dimension K dont les entrées correspondent aux K symboles transmis par l'antenne d'émission j à l'instant symbole m

Dans ces conditions nous pouvons remarquer que la puissance $p_{j,i}$ allouée au i^{ieme} symbole transmis par la j^{ieme} antenne de la station de base est la même pour tout j . L'allocation de puissance aux K utilisateurs est alors la même pour chacune des antennes de transmissions, ce qui explique que dans notre modèle la matrice des puissances \mathbf{P} ne possède pas d'indice. La matrice des puissances \mathbf{P} est alors définie par

$$\mathbf{P} = \begin{bmatrix} p_1 & & 0 \\ & \ddots & \\ 0 & & p_K \end{bmatrix}$$

et les variances des séquences transmises par chacune des antennes d'émission prennent toutes la même valeur μ donnée par

$$\mu = \frac{\sum_{i=1}^K p_i}{N} \quad (5.2)$$

5.3 Démodulation

5.3.1 Performances du filtre de Wiener sous-optimal

Afin d'estimer les symboles émis nous équipons le récepteur de chacun des K utilisateurs d'un étage d'égalisation agissant au rythme chip dont le signal de sortie est

ensuite désembrouillé et désétalé avec les codes adéquates. Le filtre de Wiener sous-optimal permettant d'estimer la séquence chip $\mathbf{d}_j(n)$ au niveau du i^{ieme} utilisateur est alors donné par

$$\mathbf{g}_{j,i} = \left(\sum_{j=1}^t \mu \mathcal{H}_{j,i} \mathcal{H}_{j,i}^H + \sigma_i^2 \mathbf{I}_r \right)^{-1} \mathbf{h}_{1,1,i} \quad (5.3)$$

avec

- $\mathbf{h}_{1,1,i}$, un vecteur de dimension $2Nr$ définie de la même façon que 3.7 à la page 46
- $\mathcal{H}_{j,i}$, une matrice de dimension $2Nr \times 3N$ définie de la même façon que 3.3 à la page 40

Première approximation du SINR

Le théorème 1 que nous avons énoncé à la page 54 nous permet d'affirmer que lorsque le nombre d'utilisateurs K et le facteur d'étalement N tendent vers l'infini dans un rapport constant β^i converge en probabilité vers $\bar{\beta}_i$ défini par

$$\bar{\beta}_i = \frac{p_i}{\mu} \frac{\int_{-1/2}^{1/2} \eta_i(e^{2i\pi f}) df}{1 - \int_{-1/2}^{1/2} \eta_i(e^{2i\pi f}) df} \quad (5.4)$$

avec $\eta_i(e^{2i\pi f})$ défini par

$$\eta_i(e^{2i\pi f}) = \mu \mathbf{h}_{1,1,i}^H(e^{2i\pi f}) [\mathbf{H}_i(e^{2i\pi f}) \mathbf{\Gamma} \mathbf{H}_i^H(e^{2i\pi f}) + \sigma_i^2 \mathbf{I}_r]^{-1} \mathbf{h}_{1,1,i}(e^{2i\pi f})$$

où $\mathbf{\Gamma}$ est une matrice diagonale de dimension $t \times t$ dont les entrées sont toutes égales à μ , soit la variance associée à chacun des flux transmis par chacune des antennes d'émission.

Seconde approximation du SINR

Le théorème 2 que nous avons énoncé à la page 58 nous permet ici d'affirmer que, lorsque le nombre d'antennes d'émission r et de réception t tendent vers l'infini dans un rapport constant, $\bar{\beta}_i$ converge en probabilité vers β_i^* défini par

$$\beta_i^* = \frac{p_i}{\mu} \left[\frac{1}{(I_t + \delta_i \mathbf{\Gamma})_{1,1}} - 1 \right] \quad (5.5)$$

où δ_i est la solution unique strictement positive de l'équation suivante

$$\delta_i = \frac{r}{t} \frac{1}{\sigma_i^2 + \frac{1}{t} \text{Trace} [\mathbf{\Gamma} (I_t + \delta_i \mathbf{\Gamma})^{-1}]} \quad (5.6)$$

La matrice diagonale $\mathbf{\Gamma}$ des variances associées au signaux transmis par les différentes antennes d'émission ayant toutes ses entrées égales à μ , les expressions de β_i^* et δ_i se simplifient de la façon suivante

$$\begin{aligned} \beta_i^* &= p_i \delta_i \\ \sigma_i^2 \delta_i + \frac{\delta_i \mu}{1 + \delta_i \mu} &= \frac{r}{t} \end{aligned} \quad (5.7)$$

Cette dernière approximation du SINR en sortie du filtre de Wiener sous-optimal est indépendante des réalisations particulières des codes de scrambling ainsi que celles du canal de propagation. Nous allons pouvoir utiliser cette expression simplifiée du SINR afin de résoudre le problème d'allocation de puissance que nous nous sommes posé au début de ce chapitre.

5.3.2 Détermination de l'allocation de puissance à long terme

Nous rappelons que résoudre notre problème d'allocation de puissance consiste à trouver le jeu de puissances p_1, \dots, p_K tel que

$$Q(\sqrt{\beta_i^*(p_1, \dots, p_K)}) = Q(\sqrt{\beta_{cible}})$$

Cela revient à déterminer p_1, \dots, p_K de sorte que

$$\beta_i^*(p_1, \dots, p_K) = \beta_{cible}$$

En injectant cette dernière expression dans 5.7 nous obtenons

$$\sigma_i^2 \frac{\beta_{cible}}{p_i} + \frac{\beta_{cible} \mu}{p_i + \beta_{cible} \mu} = \frac{r}{t} \quad (5.8)$$

Finalement, en injectant l'expression de μ donnée par 5.2, nous pouvons réécrire 5.8 sous la forme $f_i(p_1, \dots, p_K) = 0$, où

$$\begin{aligned} f_i(p_1, \dots, p_K) &= \left(\frac{\beta_{cible}}{N} - \frac{r}{t} - \frac{r}{t} \frac{\beta_{cible}}{N} \right) p_i^2 + \\ &\quad \left(\sigma_i^2 \beta_{cible} + \sigma_i^2 \frac{\beta_{cible}^2}{N} + \frac{\beta_{cible}}{N} \left(1 - \frac{r}{t} \right) \sum_{k \neq i} p_k \right) p_i + \\ &\quad \sigma_i^2 \frac{\beta_{cible}^2}{N} \sum_{k \neq i} p_k \end{aligned}$$

Nous aboutissons alors à la solution de notre problème d'allocation de puissance que nous résumons ici.

Proposition 1. *L'allocation de puissance permettant de garantir à chaque utilisateur un SINR cible β_{cible} est la solution du système d'équation non linéaire de K équations comptant K inconnues donné par*

$$f(p) = \begin{bmatrix} f_1(p_1, \dots, p_K) \\ \vdots \\ f_K(p_1, \dots, p_K) \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

où les coefficients des fonction polynomiales $f_i(p_1, \dots, p_K) = m_2 p_i^2 + m_{i,1} p_i + m_{i,0}$ sont donnés par

$$\begin{aligned} m_2 &= \frac{\beta_{cible}}{N} - \frac{r}{t} - \frac{r}{t} \frac{\beta_{cible}}{N} \\ m_{i,1} &= \sigma_i^2 \beta_{cible} + \sigma_i^2 \frac{\beta_{cible}^2}{N} + \frac{\beta_{cible}}{N} \left(1 - \frac{r}{t}\right) \sum_{k \neq i} p_k \\ m_{i,0} &= \sigma_i^2 \frac{\beta_{cible}^2}{N} \sum_{k \neq i} p_k \end{aligned}$$

La résolution de ce système d'équations nous mène alors à l'allocation de puissance (p_1, \dots, p_K) permettant de garantir à chacun des utilisateurs le SINR voulu. Une méthode de résolution directe n'étant pas possible, il est nécessaire de recourir à une méthode numérique, nous choisissons alors la méthode de Newton qui nous permet de nous approcher de la solution de façon itérative. La solution du système est alors obtenue à l'aide de la fonction récursive suivante

$$p^{(n+1)} = p^{(n)} + Jac[f(p^{(n)})]^{-1} f(p^{(n)})$$

où $p^{(n)}$ est le vecteur de dimension K des valeurs prises par les puissances à l'itération n et $Jac[f(p^{(n)})]$ la matrice jacobienne de dimension $K \times K$ du vecteur f au point $p^{(n+1)}$.

Faisabilité de la méthode dans un contexte réel

Dans l'algorithme que nous avons décrit, nous pouvons remarquer que le seul paramètre indéterminé est la variance de bruit associé au récepteur de chacun des

utilisateurs. En pratique, un lien montant devra être mis en oeuvre afin de communiquer à la station de base la valeur de σ_i^2 estimée au niveau de chacun des mobiles supportés par le réseau.

A ce sujet, il est important de remarquer qu'il est facile d'estimer la variance de bruit σ_i^2 . En effet, dans le WCDMA il existe un canal pilote véhiculant des symboles connus à puissance constante à destination de l'ensemble des utilisateurs de la station de base. Ce canal est appelé CPICH, pour Common Pilot CHannel. Il est alors facile d'estimer la variance du bruit s'ajoutant à ces symboles pilotes. Ensuite, en évaluant l'écart de puissance entre les symboles pilotes et les symboles utiles, il devient possible d'estimer σ_i^2 par une simple règle de proportionnalité.

5.4 Efficacité de la méthode d'allocation de puissance à long terme

Afin d'évaluer l'efficacité de notre solution nous décidons de la confronter à la méthode d'allocation de puissance telle qu'elle est définie dans la norme WCDMA. Nous allons donc dans un premier temps décrire cette dernière.

5.4.1 Contrôle de puissance associé au WCDMA

Le contrôle de puissance défini par 3GPP pour le WCDMA est un mécanisme à boucle fermée combinant une boucle dite interne et une autre dite externe. Le but de la boucle interne est de maintenir la valeur instantanée du SINR en sortie du filtre de réception à une valeur cible, correspondant à un taux d'erreur bloc (BLER) ou un taux d'erreur binaire (BER) donné. Le SINR cible, dont la valeur dépend du service (Voix ou données) est fixé à l'aide de la boucle externe par le RRC (Radio Resource Control) via les couches supérieures du réseau. Cette tâche est de toute première importance car elle a un impact direct sur la capacité totale du réseau.

Dans notre travail, nous fixerons arbitrairement le BER cible et nous nous intéresserons uniquement à la boucle interne dont nous détaillons ici le fonctionnement. Périodiquement, le SINR est mesuré au niveau de chacun des mobiles afin d'estimer la qualité du lien. Le mobile détermine alors si la valeur mesurée se situe en-deça ou au-delà du SINR cible que lui impose le réseau. En fonction du résultat de cette comparaison le mobile enverra sur une liaison montante dédiée une commande TPC

(Transmit Power Control) permettant à la station de base de mettre à jour le niveau de puissance allouée selon un pas de puissance fixe. En d'autres termes, ce mécanisme s'attache à adapter rapidement la puissance nécessaire à l'obtention d'un certain taux d'erreurs instantané. Plus précisément, à chacune des mesures de SINR les actions suivantes sont engagées :

- si $SINR_{est} < SINR_{cible}$, une commande $TPC = 0$ est transmise et la station de base incrémente la puissance
- si $SINR_{est} > SINR_{cible}$, une commande $TPC = 1$ est transmise et la station de base décrémente la puissance

5.4.2 Résultats de simulation

Afin de valider notre algorithme d'allocation de puissance, nous considérons une station de base équipée de $t = 4$ antennes d'émission supportant $K = 8$ utilisateurs et un facteur d'étalement fixé à $N = 32$. Les utilisateurs sont alors répartis uniformément et leurs mobiles sont équipés de $r = 4$ antennes de réception et d'un récepteur de Wiener sous-optimal. Le rayon R de la cellule ainsi formée est égal à 0.577 kilomètres et la puissance maximale d'émission est fixée à 20 Watts. Dans ces conditions, un classique bilan de liaison nous mène à considérer une valeur de BER cible égale à 10^{-2} . Les canaux de propagation sélectifs en fréquence situés entre la station de base et chacun des mobiles ont été générés selon le modèle gaussien présenté dans le chapitre 3. La longueur de la réponse impulsionnelle commune à chacun d'eux est égale à $L = 5$ et nous la considérons constante pendant la durée d'un bloc de 50 symboles.

La figure 5.1 représente

- BER_{WCDMA} , le taux d'erreur binaire moyen de l'utilisateur d'intérêt obtenu à l'aide de la méthode de contrôle de puissance de la norme WCDMA
- BER , le taux d'erreur binaire moyen de l'utilisateur d'intérêt obtenu à l'aide de notre méthode de contrôle de puissance à long terme

par rapport à la distance du mobile d'intérêt à la station de base. Plus précisément, pour chaque point de distance mobile-station de base nous avons généré aléatoirement et uniformément 100 fois la position des $K - 1$ autres utilisateurs et moyenné les taux d'erreurs binaires associés à chacune des 2 méthodes. Les SINR associés aux BER précédents ont été calculés à partir de l'équation 5.4 à l'aide d'une boucle de Monte Carlo sur 100 slots consécutifs à chacun desquels nous associons une réalisa-

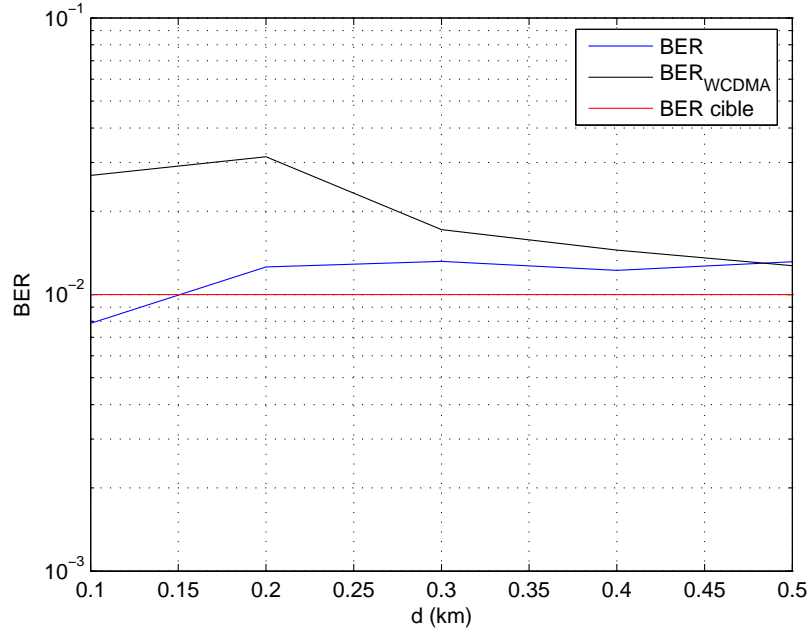


FIGURE 5.1 – BER moyen versus distance mobile-Station de base

tion du canal de propagation (canal à évanouissement par bloc). En ce qui concerne la méthode d'allocation de puissance inspirée du WCDMA, la puissance est ajustée à chacune des réalisations du canal de propagation, c'est à dire à chaque slot. Le BER de l'utilisateur d'intérêt est alors moyenné sur l'ensemble des 100 slots de chacune des 100 distributions des utilisateurs au sein de la cellule.

Tout d'abord, nous remarquons que l'allocation de puissance associée au WCDMA ne parvient jamais à satisfaire l'utilisateur. Cela s'explique de la façon suivante : tirer au sort les réalisations du canal de propagation à chaque slot revient à générer un canal de propagation entre la station de base et des utilisateurs dotés d'une très forte mobilité, se déplaçant à une vitesse de l'ordre d'une centaine de $km.h^{-1}$. Dans ces conditions, il est difficile pour la méthode associée au WCDMA de compenser une évolution si rapide des coefficients du canal de propagation, ce qui se répercute sur les performances obtenues. A l'inverse, en ce qui concerne notre méthode d'allocation de puissance, le changement rapide des coefficients du canal de propagation est un avantage car la moyenne des BER instantanés converge ainsi rapidement vers la valeur moyenne cible ayant été retenue pour le calcul des puissances à allouer à chacun des utilisateurs. Ceci explique le fait que la méthode d'allocation de puis-

sance à long terme est bien plus efficace que la méthode d'allocation de puissance du WCDMA, et ce quelque soit la distance mobile - Station de base.

Ces résultats nous assurent donc de la validité de la méthode d'allocation de puissance à long terme. On se rend compte qu'il est plus efficace de garantir un taux d'erreur moyen plutôt que de s'évertuer à maintenir le BER instantané au-dessus d'une valeur cible. Cette méthode d'allocation de puissance à long terme, en plus d'être plus précise, permet de réduire le trafic nécessaire en lien montant. En effet, la valeur de variance de bruit σ_i^2 associée au i^{ieme} utilisateur nécessaire à la mise en oeuvre de l'algorithme doit être transmise avec une période bien plus grande que celle avec laquelle il est nécessaire de remonter les indications de TPC inhérents à la mise en place du contrôle de puissance du WCDMA.

Deuxième partie

Optimisation de précodeurs

Chapitre 6

Avant-Propos

Dans la partie précédente nous nous sommes intéressés à l'étude des performances en liaison descendante du récepteur MMSE équipant les mobiles présents au sein d'un réseau cellulaire ayant recours au CDMA comme méthode d'accès multiple. Nous y avons notamment développé des approximations asymptotiques suffisamment fiables du SINR en sortie du filtre MMSE que nous avons ensuite pu utiliser afin de déterminer la solution d'un problème d'allocation de puissance ayant pour but de garantir une certaine qualité de liaison à l'ensemble des utilisateurs de la cellule.

Afin d'obtenir des résultats de nature différente et d'ordre plus général, nous avons entrepris d'étudier la conception de précodeurs visant à maximiser la capacité ergodique d'un système de transmission mono-utilisateur équipé de t antennes d'émission et r antennes de réception et du récepteur MMSE dans le cas d'un canal sélectif en fréquence. La capacité ergodique \mathcal{C}_{mmse} d'un tel système est donnée par

$$\mathcal{C}_{mmse} = -\mathbb{E} \sum_{j=1}^t \log_2 \left(\int_{-1/2}^{1/2} (\mathbf{Q}(e^{2i\pi f}))_{j,j} df \right) \quad (6.1)$$

où la matrice $\mathbf{Q}(e^{2i\pi f})$ de dimension $t \times r$ est définie par

$$\mathbf{Q}(e^{2i\pi f}) = \frac{\sigma^2}{p} \left(\mathbf{H}(e^{2i\pi f})^H \mathbf{H}(e^{2i\pi f}) + \frac{\sigma^2}{p} \mathbf{I} \right)^{-1}$$

où l'on retrouve

- $\mathbf{H}(e^{2i\pi f})$, une matrice de dimension $r \times t$ représentant la fonction de transfert fréquentielle du canal de propagation MIMO
- σ^2 , un scalaire représentant la variance du classique bruit additif gaussien au niveau du récepteur

- p , un scalaire représentant la puissance totale reçue par antenne

Nous avons choisi de nous intéresser aux cas pour lequel le canal de propagation est sujet à un phénomène de corrélation spatiale au niveau de l'émetteur. Ainsi, nous considérons un canal de propagation disposant de L trajets à chacun desquels nous associons une matrice \mathbf{H}_l de dimension $r \times t$ définie par

$$\mathbf{H}_l = p^{1/2} \frac{1}{Lt} \bar{\mathbf{H}}_l \mathbf{C}_l^{\frac{1}{2}} \quad (6.2)$$

où

- $\bar{\mathbf{H}}_l$ est une matrice gaussienne i.i.d. de dimension $r \times t$ dont les entrées sont centrées de variance unitaire
- \mathbf{C}_l est une matrice positive de dimension $t \times t$ traduisant l'impact de la corrélation entre antennes d'émission sur le l^{ime} trajet

Considérons la matrice \mathbf{C} définie par

$$\mathbf{C} = \frac{1}{L} \sum_{l=1}^L \mathbf{C}_l$$

sous l'hypothèse que $\frac{1}{t} \text{Tr} \mathbf{C} = 1$. Dans ces conditions, l'observation du fait que la matrice de covariance de $\text{vec}(\mathbf{H}(e^{2i\pi f}))$ coïncide avec $\frac{p}{t} \mathbf{C} \otimes \mathbf{I}_r$ nous permet d'écrire la matrice $\mathbf{H}(e^{2i\pi f})$ sous la forme

$$\mathbf{H}(e^{2i\pi f}) = p^{1/2} \tilde{\mathbf{H}}(e^{2i\pi f}) \mathbf{C}^{1/2} \quad (6.3)$$

où $\tilde{\mathbf{H}}(e^{2i\pi f})$ est une matrice gaussienne de dimension $r \times t$ dont les entrées sont i.i.d. centrées et de variance $\frac{1}{t}$.

L'optimisation directe de (6.1) par rapport à un éventuel précodeur est peu attractive car elle nécessite de coûteuses simulations de Monte-Carlo. Nous choisissons d'en maximiser une approximation obtenue lorsque les nombres d'antennes en émission et en réception tendent vers l'infini dans un rapport constant. En appliquant des résultats analogues à ceux de la première partie, nous obtenons une approximation $\bar{\mathcal{C}}_{mmse}$ de \mathcal{C}_{mmse} donnée par

$$\bar{\mathcal{C}}_{mmse} = \sum_{j=1}^t \log_2 \left(\frac{1}{(\mathbf{I} + \delta(\mathbf{C}) \mathbf{C})_{j,j}^{-1}} \right) \quad (6.4)$$

où δ est l'unique solution strictement positive de l'équation suivante

$$\frac{\sigma^2}{p} \delta + \frac{1}{t} \text{Tr} \delta \mathbf{C} (\mathbf{I} + \delta \mathbf{C})^{-1} = \frac{r}{t} \quad (6.5)$$

Classiquement, le transmetteur émet t séquences de symboles $((s_j(n))_{n \in \mathbb{Z}})_{j=1, \dots, t}$ mutuellement indépendantes. L'utilisation d'un précodeur, soit une matrice \mathbf{K} de dimension $t \times t$, telle que $\frac{1}{t} \text{Tr}(\mathbf{K}\mathbf{K}^H) \leq 1$, remplace le vecteur $\mathbf{s}(n) = [s_1(n), \dots, s_t(n)]^T$ des symboles émis par le transmetteur par le vecteur $\mathbf{K}\mathbf{s}(n)$. L'impact d'un précodeur \mathbf{K} sur l'approximant de la capacité ergodique se déduit très simplement en remplaçant dans les expressions 6.4 et 6.5 la matrice de covariance \mathbf{C} par la matrice $\mathbf{K}^H \mathbf{C} \mathbf{K}$. Dans ces conditions, l'approximation de la capacité ergodique ainsi que le paramètre δ deviennent des fonction de \mathbf{K} que l'on note respectivement $\bar{\mathcal{C}}_{mmse}(\mathbf{K})$ et $\delta(\mathbf{K})$. En considérant $\mathbf{C} = \mathbf{U}\mathbf{D}\mathbf{U}^H$, la décomposition en valeur et vecteurs propres de la matrice \mathbf{C} , nous montrons qu'un précodeur optimal est donné par

$$\mathbf{K}_{opt} = \mathbf{U}\mathbf{D}^{-1/2}\mathbf{\Lambda}_{opt}^{1/2} \quad (6.6)$$

où $\mathbf{\Lambda}$ est une matrice diagonale positive dont les entrées répondent à la contrainte

$$\mathbf{\Lambda} = \text{Diag}(\lambda_1, \dots, \lambda_t) \geq 0, \quad \frac{1}{t} \text{Tr}(\mathbf{D}^{-1}\mathbf{\Lambda}) \leq 1 \quad (6.7)$$

et sont solutions d'un certain problème d'optimisation.

Ces résultats ont ainsi pu faire l'objet d'un article de conférence intitulé *On the ergodic capacity of frequency selective MIMO systems equipped with MMSE receivers* présenté à la conférence *IEEE Globecom 2008* que le lecteur pourra trouver au chapitre 7. Nous y présentons notamment des résultats de simulation montrant l'efficacité de précodeurs conçus sur le modèle cité précédemment, lorsque le nombre de trajets du canal de propagation est suffisamment grand. En revanche, nous montrons que l'erreur relative d'approximation de la capacité ergodique réelle par $\bar{\mathcal{C}}_{mmse}$, un terme en $O(\frac{1}{t})$, devient trop importante lorsque le canal de propagation ne compte qu'un seul trajet, c'est à dire dans le cas des canaux plats.

L'étude d'un approximant en grande dimension des performances du récepteur MMSE dans le cadre de canaux MIMO plats i.i.d. a déjà été rigoureusement étudié par [18] à l'aide de méthodes basées sur les matrices aléatoires. Les auteurs montrent que, lorsque le SNR est fixé et que les nombres d'antennes en émission et en réception tendent vers l'infini, la distribution limite de la capacité ergodique associée au récepteur MMSE est gaussienne et que son espérance et sa variance peuvent être exprimée algébriquement. Le terme d'erreur associé à l'approximant ainsi proposé est alors un terme en $O(\frac{1}{t})$. Les performances asymptotique du récepteur MMSE dans le contexte des canaux MIMO plats doublement corrélés suivant un modèle de

Kronecker ont été étudiées par [17] à l'aide de la méthode des répliques, dont la pertinence mathématique reste à démontrer. Les auteurs montrent dans ce contexte que la capacité ergodique associée au récepteur MMSE converge en distribution vers une variable aléatoire gaussienne dont ils calculent l'espérance et la variance. Ici, nous montrons rigoureusement comment étudier le comportement en grande dimension de la capacité ergodique dans le cadre de canaux MIMO plats doublement corrélés à l'aide de matrices aléatoires et déterminons ainsi une approximation de la capacité ergodique dont l'erreur relative associée est un terme en $O\left(\frac{1}{t^2}\right)$. Nos résultats de simulations nous assurent de la validité de cette approximation pour un faible nombre d'antennes d'émission et de réception.

Nous considérons donc un canal de propagation plat doublement corrélé et lui associons une matrice aléatoire gaussienne \mathbf{H} définie par

$$\mathbf{H} = \frac{1}{\sqrt{t}} \mathbf{C}_R^{1/2} \mathbf{H}_{iid} \mathbf{C}_T^{1/2} \quad (6.8)$$

où

- \mathbf{H}_{iid} est une matrice gaussienne de dimension $r \times t$ dont les entrées sont des variables aléatoires complexes circulaires i.i.d. centrées de variance unitaire
- \mathbf{C}_T est une matrice positive de dimension $t \times t$ traduisant le phénomène de corrélation au niveau du réseau d'antennes d'émission
- \mathbf{C}_R est une matrice positive de dimension $r \times r$ traduisant le phénomène de corrélation au niveau du réseau d'antennes de réception

Dans ces conditions l'expression de $\mathcal{C}_{mmse}(\mathbf{K})$, la capacité ergodique dépendant du précodeur \mathbf{K} , est donnée par

$$\mathcal{C}_{mmse}(\mathbf{K}) = -\mathbb{E} \left[\sum_{j=1}^t \log (\sigma^2 \mathbf{Q}_T(\mathbf{K}))_{j,j} \right] \quad (6.9)$$

où la matrice $\mathbf{Q}_T(\mathbf{K})$ est elle-même donnée par

$$\mathbf{Q}_T(\mathbf{K}) = (\mathbf{K}^H \mathbf{H}^H \mathbf{H} \mathbf{K} + \sigma^2 \mathbf{I}_t)^{-1} \quad (6.10)$$

De la même façon que précédemment, nous devons déterminer un approximant en grande dimension de $\mathcal{C}_{mmse}(\mathbf{K})$ que l'on utilisera dans le but de simplifier le problème de maximisation de la capacité ergodique. Une application des résultats précédents ainsi qu'un développement au second ordre de $\sigma^2 \log \mathbf{Q}_{T,j,j}(\mathbf{K})$ autour de $\mathbb{E}(\mathbf{Q}_{T,j,j}(\mathbf{K}))$ nous mène alors à une nouvelle approximation $\hat{\mathcal{C}}_{mmse}(\mathbf{K})$, dont l'erreur

relative associée est cette fois-ci un terme en $O\left(\frac{1}{t^2}\right)$. Si nous considérons la matrice $\mathbf{T}_T(\mathbf{K})$ donnée par

$$\mathbf{T}_T(\mathbf{K}) = \left(\sigma^2 (\mathbf{I} + \delta(\mathbf{K}) \mathbf{K}^H \mathbf{C}_T \mathbf{K})\right)^{-1}$$

alors $\hat{\mathcal{C}}_{mmse}(\mathbf{K})$ s'exprime de la façon suivante

$$\hat{\mathcal{C}}_{mmse}(\mathbf{K}) = \bar{\mathcal{C}}_{mmse}(\mathbf{K}) + \frac{1}{2\tilde{\delta}(\mathbf{K})^2} \frac{\tilde{\gamma}(\mathbf{K})}{1 - \sigma^4 \gamma(\mathbf{K}) \tilde{\gamma}(\mathbf{K})} \frac{1}{t} \sum_{j=1}^t \left(1 - \frac{\left((\sigma^2 \mathbf{T}_T(\mathbf{K}))^2\right)_{j,j}}{\sigma^2 \mathbf{T}_{T,j,j}(\mathbf{K})} \right)^2 \quad (6.11)$$

où l'on retrouve

– $\bar{\mathcal{C}}_{mmse}(\mathbf{K})$, l'approximant classique de la capacité ergodique, donné par

$$\bar{\mathcal{C}}_{mmse}(\mathbf{K}) = \sum_{j=1}^t \log_2 \left(\frac{1}{(\mathbf{I} + \delta(\mathbf{K}) \mathbf{K}^H \mathbf{C}_T \mathbf{K})_{j,j}^{-1}} \right) \quad (6.12)$$

où $\delta(\mathbf{K})$ est la première composante de l'unique couple solution du système d'équations suivant

$$\delta(\mathbf{K}) = \frac{1}{t} \text{Tr} \left[\mathbf{K}^H \mathbf{C}_T \mathbf{K} \left(\sigma^2 (\mathbf{I} + \tilde{\delta}(\mathbf{K}) \mathbf{K}^H \mathbf{C}_T \mathbf{K}) \right)^{-1} \right] \quad (6.13)$$

$$\tilde{\delta}(\mathbf{K}) = \frac{1}{t} \text{Tr} \left[\mathbf{C}_R (\sigma^2 (\mathbf{I} + \delta(\mathbf{K}) \mathbf{C}_R))^{-1} \right] \quad (6.14)$$

$$- \gamma(\mathbf{K}) = \frac{1}{t} \text{Tr} \left[(\mathbf{K}^H \mathbf{C}_T \mathbf{K})^2 \left(\sigma^2 (\mathbf{I} + \tilde{\delta}(\mathbf{K}) \mathbf{K}^H \mathbf{C}_T \mathbf{K}) \right)^{-2} \right]$$

$$- \tilde{\gamma}(\mathbf{K}) = \frac{1}{t} \text{Tr} \left[\mathbf{C}_R^2 (\sigma^2 (\mathbf{I} + \delta(\mathbf{K}) \mathbf{C}_R))^{-2} \right]$$

Nous montrons qu'un précodeur \mathbf{K} tel que $\frac{1}{t} \text{Tr}(\mathbf{K} \mathbf{K}^H) \leq 1$ et maximisant $\hat{\mathcal{C}}_{mmse}(\mathbf{K})$ diagonalise la matrice de covariance à l'émetteur $\mathbf{K}^H \mathbf{C}_T \mathbf{K}$ et nous définissons une structure de précodeur optimal analogue à celle donnée précédemment, soit

$$\mathbf{K}_{opt} = \mathbf{U} \mathbf{D}^{-1/2} \mathbf{\Lambda}_{opt}^{1/2} \quad (6.15)$$

où $\mathbf{\Lambda}$ est une matrice diagonale positive dont les entrées répondent à la contrainte

$$\mathbf{\Lambda} = \text{Diag}(\lambda_1, \dots, \lambda_t) \geq 0, \quad \frac{1}{t} \text{Tr}(\mathbf{D}^{-1} \mathbf{\Lambda}) \leq 1 \quad (6.16)$$

et sont solutions d'un nouveau problème d'optimisation.

Ce travail a ainsi pu faire l'objet d'un article, que le lecteur pourra trouver au chapitre 8, intitulé *On the precoder design of flat fading MIMO systems equipped with*

MMSE receivers : a large system approach que nous avons soumis à la revue *IEEE Transaction on Information Theory*. Les résultats de simulation qui y sont présentés nous assurent de la validité de notre structure de précodeurs dans le contexte des canaux plats doublement corrélés.

Chapitre 7

Article présenté à IEEE Globecom 2008

Sommaire

7.1	Abstract	99
7.2	Introduction	99
7.3	Asymptotic expression of \mathcal{C}_{mmse}	100
7.4	Design of optimal precoders.	103
7.5	Concluding remarks.	108

On the ergodic capacity of frequency selective MIMO systems equipped with MMSE receivers.

Cedric Artigue, Philippe Loubaton

7.1 Abstract

This paper is devoted to the study of the ergodic capacity of frequency selective MIMO systems equipped with a MMSE receiver when the channel state information is available at the receiver side and when the second order statistics of the channel taps are known at the transmitter side. As its expression is rather complicated and difficult to analyse, it is studied in the case where the number of transmit and receive antennas converge to $+\infty$ at the same rate. In this asymptotic regime, the main results of the paper are related to the design of an optimal precoder in the case where the transmit antennas are correlated. It is shown that the left singular eigenvectors of the optimum precoder coincide with the eigenvectors of the mean of the channel taps transmit covariance matrices, and its singular values are solution of a certain minimization problem.

7.2 Introduction

It is now well established that using multiple transmit and receive antennas potentially allows to increase the Shannon capacity of digital communications systems. Since the seminal work of Teletar ([1]), the ergodic Shannon capacity of block fading MIMO systems has been studied extensively. If the channel state information is available at the receiver side while the transmitter is only aware of its second order statistical properties, important questions such as the impact of channel correlations on the capacity or the design of optimal precoding schemes have been addressed by several authors (see e.g. [31] for a review).

The ergodic Shannon capacity is certainly a valuable figure of merit if the MIMO system under consideration is equipped with a maximum likelihood decoder. As this receiver may be difficult to use in practice, in particular if the MIMO channel is frequency selective, it is also quite relevant to study the potential performance of MIMO systems equipped with suboptimum linear interfaces. In this paper, we consider the MMSE receiver, and study the corresponding ergodic capacity denoted \mathcal{C}_{mmse} in the following. It is defined as the sum over the transmit antennas of the terms $\mathbb{E}(\log(1 + \text{SINR}_j))$, where SINR_j represents the SINR associated to the stream sent by antenna j provided by the MMSE receiver. \mathcal{C}_{mmse} has been studied extensively in the past if the channel is static and available at the transmitter; in this context, the design of an optimum precoder has been addressed in a number of papers (see e.g. [8]). However, assuming that the transmitter is aware of the channel state information is often not realistic in the context of mobile wireless systems. In this paper, we assume that the MIMO channel is frequency selective, known at the receiver side, but that its second order statistics of the channel are available at the transmitter side. As in [9], the channel taps are modelled as independent Gaussian random matrices with possible transmit correlations. To our best knowledge, \mathcal{C}_{mmse} seems not to have been studied extensively in the past.

As the expression of \mathcal{C}_{mmse} is a rather complicated, we evaluate its behaviour in the case where the number of transmit and receive antennas converge to $+\infty$ at the same rate. In this asymptotic regime, \mathcal{C}_{mmse} has the same behaviour that a simpler term $\bar{\mathcal{C}}_{mmse}$. As observed in a number of situations, $\bar{\mathcal{C}}_{mmse}$ is a reliable approximation of \mathcal{C}_{mmse} , even for a realistic number of transmit and receive antennas. We show that the transmission scheme for which $\bar{\mathcal{C}}_{mmse}$ is optimum is a (non frequency dependent) linear precoding. We address the optimization of $\bar{\mathcal{C}}_{mmse}$ w.r.t. the precoder, and show that the left singular vectors of the optimum precoder(s) coincide with the eigenvectors of a covariance matrix depending on second order statistics of the channel taps. The singular values are moreover solution of a certain optimization problem.

7.3 Asymptotic expression of \mathcal{C}_{mmse} .

We consider a MIMO system equipped with r receive antennas and t transmit antennas. The transfer function $\mathbf{H}(z)$ of discrete-time equivalent MIMO channel

is given by $\mathbf{H}(z) = \sum_{l=0}^{L-1} \mathbf{H}_l z^{-l}$. The channel taps $(\mathbf{H}_l)_{l=0,\dots,L-1}$ are assumed to be mutually independent Gaussian random matrices. As in [9]¹, each matrix \mathbf{H}_l is modelled as

$$\mathbf{H}_l = p^{1/2} \frac{1}{\sqrt{Lt}} \bar{\mathbf{H}}_l \mathbf{C}_l^{1/2} \quad (7.1)$$

where p represents the total received power per receive antenna and where $\bar{\mathbf{H}}_l$ is a zero mean complex Gaussian matrix with unit variance independent identically distributed (i.i.d.) entries. Matrix \mathbf{C}_l is a positive matrix modelling the impact of transmit antenna correlation on the l -th path. $p \frac{1}{L} \frac{1}{t} \text{Tr}(\mathbf{C}_l)$ represents the power of path l . If we denote \mathbf{C} the matrix

$$\mathbf{C} = \frac{1}{L} \sum_{l=0}^{L-1} (\mathbf{C}_l)$$

we assume that $\frac{1}{t} \text{Tr} \mathbf{C} = 1$. These normalizations ensure that the total received power is independent of L and t and thus allows to compare in a fair fashion situations in which the number of paths and the number of transmit antennas differ. Model (7.1) implies that $\mathbf{H}(e^{2i\pi f})$ can be written as

$$\mathbf{H}(e^{2i\pi f}) = p^{1/2} \tilde{\mathbf{H}}(e^{2i\pi f}) \mathbf{C}^{1/2} \quad (7.2)$$

where matrix $\tilde{\mathbf{H}}(e^{2i\pi f})$ is a zero mean Gaussian $r \times t$ random matrix with variance $\frac{1}{t}$ i.i.d. entries. This is based on the observation that the covariance matrix of $\text{vec}(\mathbf{H}(e^{2i\pi f}))$ coincides with $p/t \mathbf{C} \otimes \mathbf{I}_r$ (see [9] for more details). This means that for each f , $\mathbf{H}(e^{2i\pi f})$ can be interpreted as the channel matrix of a semi-correlated flat fading MIMO channel whose transmit correlation matrix is independent of the frequency f .

The transmitter sends t streams of symbols $((s_j(n))_{n \in \mathbb{Z}})_{j=1,\dots,t}$ (one stream per transmit antenna). In this section, we assume that the $(s_j)_{j=1,\dots,t}$ are unit variance mutually independent sequences. The corresponding r -variate discrete-time received signal $(\mathbf{y}(n))_{n \in \mathbb{Z}}$ is given by

$$\mathbf{y}(n) = \sum_{l=0}^{L-1} \mathbf{H}_l \mathbf{s}(n-l) + \mathbf{v}(n)$$

where $\mathbf{s}(n) = (s_1(n), \dots, s_t(n))^T$ and where \mathbf{v} is a white Gaussian noise with covariance matrix $\mathbb{E}(\mathbf{v}(n)\mathbf{v}(n)^H) = \sigma^2 \mathbf{I}_r$. Each symbol sequence s_j is estimated by the

1. [9] addresses uplink transmissions so that the receive antennas are correlated. Here, we rather focus here on downlink transmissions and the transmit antennas are correlated

non causal Wiener filter whose transfer function is

$$\mathbf{h}_j(e^{2i\pi f})^H (\mathbf{H}(e^{2i\pi f}) \mathbf{H}(e^{2i\pi f})^H + \sigma^2 \mathbf{I}_r)^{-1}$$

where $\mathbf{h}_j(e^{2i\pi f})$ is the j -th column of $\mathbf{H}(e^{2i\pi f})$. It is standard that the SINR β_j provided by this linear receiver is given by

$$\beta_j = \frac{1}{\int_{-1/2}^{1/2} (\mathbf{Q}(e^{2i\pi f}))_{j,j} df} - 1 \quad (7.3)$$

where matrix $\mathbf{Q}(e^{2i\pi f})$ is defined by

$$\mathbf{Q}(e^{2i\pi f}) = \frac{\sigma^2}{p} \left(\mathbf{H}(e^{2i\pi f})^H \mathbf{H}(e^{2i\pi f}) + \frac{\sigma^2}{p} \mathbf{I} \right)^{-1}$$

The ergodic capacity \mathcal{C}_{mmse} of the MIMO system under consideration is thus equal to

$$\mathbb{E} \sum_{j=1}^t (\log_2(1 + \beta_j)) = -\mathbb{E} \sum_{j=1}^t \log_2 \left(\int_{-1/2}^{1/2} (\mathbf{Q}(e^{2i\pi f}))_{j,j} df \right) \quad (7.4)$$

where the mathematical expectation is over the joint probability distribution of random matrices $\mathbf{H}_0, \dots, \mathbf{H}_{L-1}$. It is clear the expression of \mathcal{C}_{mmse} is complicated and difficult to exploit.

When r and t converge to $+\infty$ at the same rate, the entries of $\mathbf{Q}(e^{2i\pi f})$ converge towards deterministic terms which only depend on the statistical properties of $\mathbf{H}(e^{2i\pi f})$. By (7.2), the second order statistical properties of $\mathbf{H}(e^{2i\pi f})$ are independent of the frequency f . Therefore, the large system approximation of the entries $\mathbf{Q}(e^{2i\pi f})$ are independent of f . More precisely, it is possible to prove the following theorem.

Theorem 1. *For each f , when t and r converge to $+\infty$ in such a way that $\frac{r}{t} \rightarrow \alpha$, then, the entries of matrix $\mathbf{Q}(e^{2i\pi f})$ converge in probability towards the entries of the deterministic matrix $(\mathbf{I}_t + \delta(\mathbf{C}) \mathbf{C})^{-1}$, where $\delta(\mathbf{C})$ is the unique strictly positive solution of the equation*

$$\frac{\sigma^2}{p} \delta + \frac{1}{t} \text{Tr} \delta \mathbf{C} (\mathbf{I} + \delta \mathbf{C})^{-1} = \frac{r}{t} \quad (7.5)$$

Using Theorem 1, we get immediately that \mathcal{C}_{mmse} has the same asymptotic behaviour that $\bar{\mathcal{C}}_{mmse}$ defined by

$$\bar{\mathcal{C}}_{mmse} = \sum_{j=1}^t \log_2 \left(\frac{1}{(\mathbf{I} + \delta(\mathbf{C}) \mathbf{C})_{j,j}^{-1}} \right) \quad (7.6)$$

In other words, the relative error $\frac{\bar{\mathcal{C}}_{mmse} - \mathcal{C}_{mmse}}{\mathcal{C}_{mmse}}$ converges towards 0 in the above asymptotic regime.

We now present some simulation experiments which show that the number of antennas r and t for which the approximation $\bar{\mathcal{C}}_{mmse} \simeq \mathcal{C}_{mmse}$ is relevant depends on a crucial way on the number of paths L . In Figure 7.1, we represent the average of the absolute value of the relative error between $\bar{\mathcal{C}}_{mmse}$ and \mathcal{C}_{mmse} when $L = 1, 4, 7, 10$ for $r = t = 4$, $\frac{p}{\sigma^2} = 10$ dB, and $\mathbf{C} = \mathbf{I}$. The average of the relative error is evaluated on 200 independent realizations of the matrices $(\mathbf{H}(l))_{l=0,\dots,L-1}$. It is clear that if $L = 1$, then the large system approximant is irrelevant, but that for $L = 4, 7, 10$, the error becomes reasonable.

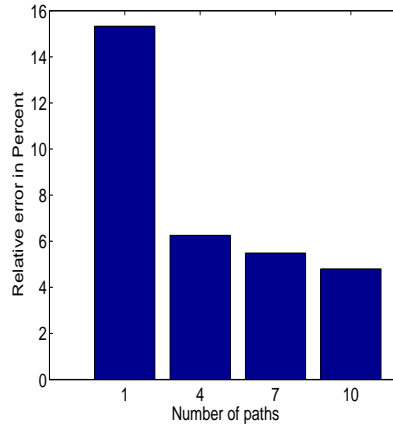


FIGURE 7.1 – Average of the relative error.

The above discussion shows that in rich scattering environments, our large system approximation of \mathcal{C}_{mmse} is reliable for realistic numbers of transmit and receive antennas. Therefore, it is relevant to study $\bar{\mathcal{C}}_{mmse}$ in place of \mathcal{C}_{mmse} . In the rest of the paper, we study the effect of \mathbf{C} on $\bar{\mathcal{C}}_{mmse}$, and discuss on the design of linear precoders that optimize $\bar{\mathcal{C}}_{mmse}$.

7.4 Design of optimal precoders.

7.4.1 Influence of the transmit correlation matrix \mathbf{C} on $\bar{\mathcal{C}}_{mmse}$.

In order to motivate the use of linear precoders, we first establish in this paragraph that diagonal correlation matrices provide higher values of $\bar{\mathcal{C}}_{mmse}$. In the

remainder of this paragraph, we mention explicitly that $\bar{\mathcal{C}}_{mmse}$ depends on $\mathbf{C} = \frac{1}{L} \sum_{l=0}^{L-1} \mathbf{C}_l$ and use the notation $\bar{\mathcal{C}}_{mmse}(\mathbf{C})$. We have the following result.

Proposition 1. *Let \mathbf{C} be a positive matrix satisfying $\frac{1}{t} \text{Tr}(\mathbf{C}) = 1$, and let $\mathbf{C} = \mathbf{U}\mathbf{D}\mathbf{U}^H$ be its eigenvalues/ eigenvectors decomposition, where the diagonal entries $(d_j)_{j=1,\dots,t}$ of \mathbf{D} are arranged in decreasing order. Then,*

$$\bar{\mathcal{C}}_{mmse}(\mathbf{C}) \leq \bar{\mathcal{C}}_{mmse}(\mathbf{D}) \quad (7.7)$$

Proof. We first note that parameter δ defined by (7.5) depends on matrix \mathbf{C} , but through its eigenvalues. Therefore, $\delta(\mathbf{C})$ coincides with $\delta(\mathbf{D})$. $\bar{\mathcal{C}}_{mmse}(\mathbf{C})$ can thus be written as

$$\bar{\mathcal{C}}_{mmse}(\mathbf{C}) = \sum_{j=1}^t \log_2 \left(\frac{1}{(\mathbf{I} + \delta(\mathbf{D})\mathbf{U}\mathbf{D}\mathbf{U}^H)_{j,j}^{-1}} \right)$$

$(\mathbf{I} + \delta(\mathbf{D})\mathbf{U}\mathbf{D}\mathbf{U}^H)_{j,j}^{-1}$ is given by

$$(\mathbf{I} + \delta(\mathbf{D})\mathbf{U}\mathbf{D}\mathbf{U}^H)_{j,j}^{-1} = \sum_{k=1}^t \frac{|u_{k,j}|^2}{1 + \delta(\mathbf{D})d_k}$$

where $u_{k,j}$ is the entry (k, j) of unitary matrix \mathbf{U} . The function $y \rightarrow \log_2 \frac{1}{y}$ is convex on \mathbb{R}^+ . As $\sum_{k=1}^t |u_{k,j}|^2 = 1$ (because \mathbf{U} is unitary), we have

$$\log_2 \left[\frac{1}{\sum_{k=1}^t \frac{|u_{k,j}|^2}{1 + \delta(\mathbf{D})d_k}} \right] \leq \sum_{k=1}^t |u_{k,j}|^2 \log_2(1 + \delta(\mathbf{D})d_k)$$

Summing over j , and using that $\sum_{k=1}^t |u_{k,j}|^2 = 1$, we get that

$$\bar{\mathcal{C}}_{mmse}(\mathbf{C}) \leq \sum_{k=1}^t \log_2(1 + \delta(\mathbf{D})d_k) = \bar{\mathcal{C}}_{mmse}(\mathbf{D})$$

This result implies that the maximum of $\bar{\mathcal{C}}_{mmse}(\mathbf{C})$ over the set of all possible correlation matrices for which $\frac{1}{t} \text{Tr}(\mathbf{C}) = 1$ is reached for a diagonal correlation matrix.

7.4.2 Structure of linear precoders.

Proposition 1 suggests that the use of a linear precoder at the transmitter side may be fruitful. Let \mathbf{K} be a $t \times t$ matrix satisfying $\frac{1}{t} \text{Tr}(\mathbf{K}\mathbf{K}^H) = 1$, and assume that vector $\mathbf{K}\mathbf{s}_n$ is transmitted instead of vector \mathbf{s}_n at each time n . The capacity \mathcal{C}_{mmse} of this precoded system equipped with the MMSE receiver is of course obtained by

replacing covariance matrix \mathbf{C} by matrix $\mathbf{K}^H \mathbf{C} \mathbf{K}$. More precisely, \mathcal{C}_{mmse} is given by (8.5), but in which matrix $\mathbf{Q}(e^{2i\pi f})$ is now defined by

$$\mathbf{Q}(e^{2i\pi f}) = \frac{\sigma^2}{p} \left(\mathbf{K}^H \mathbf{C}^{1/2} \tilde{\mathbf{H}}(e^{2i\pi f})^H \tilde{\mathbf{H}}(e^{2i\pi f}) \mathbf{C}^{1/2} \mathbf{K} + \frac{\sigma^2}{p} \right)_{j,j}^{-1}$$

where $\tilde{\mathbf{H}}(e^{2i\pi f})$ is a zero mean i.i.d. Gaussian random matrix with $\frac{1}{t}$ variance entries. In order to design the precoder \mathbf{K} , it is quite natural to optimize \mathcal{C}_{mmse} over the set of all precoding matrices \mathbf{K} for which $\frac{1}{t} \text{Tr}(\mathbf{K} \mathbf{K}^H) = 1$. As the expression of \mathcal{C}_{mmse} is quite complicated, we propose to study the maximization of the large system approximation $\bar{\mathcal{C}}_{mmse}$ versus \mathbf{K} . $\bar{\mathcal{C}}_{mmse}(\mathbf{K})$ is now given by

$$\bar{\mathcal{C}}_{mmse}(\mathbf{K}) = \sum_{j=1}^t \log_2 \left(\frac{1}{(\mathbf{I} + \delta(\mathbf{K}^H \mathbf{C} \mathbf{K}) \mathbf{K}^H \mathbf{C} \mathbf{K})_{j,j}^{-1}} \right) \quad (7.8)$$

where $\delta(\mathbf{K}^H \mathbf{C} \mathbf{K})$ is the unique positive solution of the equation (7.5) in which matrix \mathbf{C} is replaced by matrix $\mathbf{K}^H \mathbf{C} \mathbf{K}$. The main result of this paper is the following theorem.

Theorem 2. *Let $\mathbf{C} = \mathbf{U} \mathbf{D} \mathbf{U}^H$ be the eigenvalue/eigenvector decomposition of matrix \mathbf{C} . Let \mathbf{K}_{opt} be the precoding matrix \mathbf{K}_{opt} defined by*

$$\mathbf{K}_{opt} = \mathbf{U} \mathbf{D}^{-1/2} \mathbf{\Gamma}_{opt}^{1/2} \quad (7.9)$$

where $\mathbf{\Gamma}_{opt} = \text{Diag}(\gamma_{1,opt}, \dots, \gamma_{t,opt})$ is a positive diagonal matrix solution of the optimization problem

Problem 1. *Maximize $\sum_{j=1}^t \log_2(1 + \gamma_j \delta(\mathbf{\Gamma}))$ under the constraints*

$$\mathbf{\Gamma} = \text{Diag}(\gamma_1, \dots, \gamma_t) \geq 0, \quad \frac{1}{t} \text{Tr}(\mathbf{D}^{-1} \mathbf{\Gamma}) = 1 \quad (7.10)$$

where $\delta(\mathbf{\Gamma})$ is the unique positive solution of the equation (7.5) in which \mathbf{C} is replaced by $\mathbf{\Gamma}$.

Proof. In order to prove Theorem 4, we consider a precoding matrix \mathbf{K} such that $\frac{1}{t} \text{Tr}(\mathbf{K} \mathbf{K}^H) = 1$, and the eigenvalue/ eigenvector decomposition $\mathbf{K}^H \mathbf{C} \mathbf{K} = \mathbf{W} \mathbf{\Gamma} \mathbf{W}^H$ of matrix $\mathbf{K}^H \mathbf{C} \mathbf{K}$. Then, the precoding matrix $\mathbf{K}_d = \mathbf{K} \mathbf{W}$ satisfies $\frac{1}{t} \text{Tr}(\mathbf{K}_d \mathbf{K}_d^H) = 1$ and is such that $\mathbf{K}_d^H \mathbf{C} \mathbf{K}_d$ is diagonal. More importantly, we have

$$\bar{\mathcal{C}}_{mmse}(\mathbf{K}) \leq \bar{\mathcal{C}}_{mmse}(\mathbf{K}_d) \quad (7.11)$$

(7.11) is shown as Proposition 1. This discussion implies that it is sufficient to look for precoding matrices \mathbf{K} for which $\mathbf{K}^H \mathbf{C} \mathbf{K}$ is a positive diagonal matrix $\mathbf{\Gamma}$. This condition is satisfied if and only if \mathbf{K} can be written as $\mathbf{K} = \mathbf{U} \mathbf{D}^{-1/2} \mathbf{\Theta} \mathbf{\Gamma}^{1/2}$ where $\mathbf{\Theta}$ is a unitary matrix. As $\frac{1}{t} \text{Tr}(\mathbf{K} \mathbf{K}^H)$ is supposed to be equal to 1, matrices $\mathbf{\Gamma}$ and $\mathbf{\Theta}$ moreover satisfy $\frac{1}{t} \text{Tr} \mathbf{D}^{-1} \mathbf{\Theta} \mathbf{\Gamma} \mathbf{\Theta}^H = 1$. Each precoder \mathbf{K} defined above can thus be parameterized by the unitary matrix $\mathbf{\Theta}$ and the positive diagonal matrix $\mathbf{\Gamma}$, or equivalently by the hermitian matrix $\mathbf{R} = \mathbf{\Theta} \mathbf{\Gamma} \mathbf{\Theta}^H$ because matrix $\mathbf{\Theta} \mathbf{\Gamma}^{1/2}$ is uniquely defined from \mathbf{R} . $\mathcal{C}_{mmse}(\mathbf{K})$ is equal to $\sum_{j=1}^t \log_2(1 + \delta(\mathbf{\Gamma}) \gamma_j)$, or equivalently to $\log_2 \det(\mathbf{I} + \delta(\mathbf{R}) \mathbf{R})$ because the eigenvalues of \mathbf{R} coincide with the entries of matrix $\mathbf{\Gamma}$, which in particular imply that $\delta(\mathbf{\Gamma}) = \delta(\mathbf{R})$. Hence, the optimization of $\bar{\mathcal{C}}_{mmse}(K)$ is equivalent to the following problem

Problem 2. *Maximize $\log_2 \det(\mathbf{I} + \delta(\mathbf{R}) \mathbf{R})$ under the constraints*

$$\mathbf{R} \geq 0, \quad \frac{1}{t} \text{Tr} \mathbf{D}^{-1} \mathbf{R} = 1, \quad (7.12)$$

In order to complete the proof of Theorem 4, it is sufficient to prove that diagonal matrices are solutions of Problem 2. This is established in the appendix.

7.4.3 Study of Problem 1.

Theorem 4 shows that the determination of an optimal precoder \mathbf{K}_{opt} needs to solve the optimization Problem 1. This problem cannot be solved in closed form, except if $\mathbf{D} = \mathbf{I}$. In this context, the results of [7] imply that the solutions of Problem 1 are the vectors $\gamma_{opt} = (\gamma_{1,opt}, \dots, \gamma_{t,opt})$ whose s largest components coincide with $\frac{t}{s}$, and whose $t - s$ smallest one are 0. The value of $s \leq t$ depends on the signal to noise ratio. If $s < t$, Problem 1 has therefore multiple solutions. This in particular implies that for $\mathbf{D} = \mathbf{I}$, the function

$$\sum_{j=1}^t \log_2(1 + \gamma_j \delta(\mathbf{\Lambda}))$$

is not concave on the convex set defined by the constraints (8.44).

If $\mathbf{D} \neq \mathbf{I}$, it seems difficult to characterize analytically the solutions of Problem 1. We have thus to use numerical technics. We propose to parameterize γ_j by $\gamma_j = \beta_j^2$ in order to get rid of the constraint $\gamma_j \geq 0$, and to use a standard gradient algorithm with projection on the constraint

$$\frac{1}{t} \sum_{j=1}^t \frac{\beta_j^2}{d_j} = 1$$

at each iteration. Note that the convergence of this algorithm towards a global maximum of $\sum_{j=1}^t \log_2(1 + \gamma_j \delta(\mathbf{\Lambda}))$ is not guaranteed because, as explained above, this last function is in general not concave. However, we have checked if $\mathbf{D} = \mathbf{I}$, then the gradient algorithm always converges towards one of the above mentioned global maximum.

7.4.4 Numerical illustration of the precoder optimization.

In order to illustrate the impact of the precoder optimization on $\bar{\mathcal{C}}_{mmse}$ and \mathcal{C}_{mmse} in a realistic context, we follow the propagation model introduced in [9] in which each path corresponds to a scatterer cluster characterized by a mean angle of departure and an angle spread. We refer the reader to [9] for more details on the corresponding correlation matrix.

In the following numerical experiment, $r = t = 4$, the number of paths is $L = 5$, and their mean departure angles and angle spreads are equal to. In Figure 7.2, we have represented $\bar{\mathcal{C}}_{mmse}(\mathbf{I})$ (without precoding) as well as $\bar{\mathcal{C}}_{mmse}(\mathbf{K}_{opt})$ where \mathbf{K}_{opt} is the optimal precoder. We have of course also represented the true ergodic capacities $\mathcal{C}_{mmse}(\mathbf{I})$ and $\mathcal{C}_{mmse}(\mathbf{K}_{opt})$ evaluated by Monte Carlo simulations. Figure 7.2 confirms that \mathcal{C}_{mmse} and $\bar{\mathcal{C}}_{mmse}$ are reasonably close one from each other. In particular, optimizing $\bar{\mathcal{C}}_{mmse}$ over the precoder also allows to improve a lot the true capacity \mathcal{C}_{mmse} .

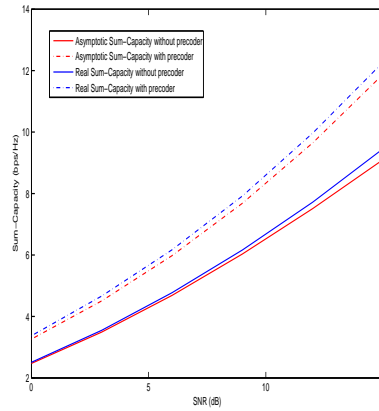


FIGURE 7.2 – Capacity with and without optimum precoding.

7.5 Concluding remarks.

We finally summarize the advantages of our asymptotic analysis of \mathcal{C}_{mmse} . It first allows to prove the relevance of precoders $\mathbf{K} = \mathbf{U}\mathbf{D}^{-1/2}\mathbf{\Gamma}^{1/2}$, where $\mathbf{\Gamma}$ is a positive diagonal matrix. Second, the entries of the optimum matrix $\mathbf{\Gamma}$ are solution of an optimization problem that can be solved by a computationnally attractive gradient algorithm. It is important to remark that it could also be possible to look for the matrix $\mathbf{\Gamma}$ maximizing w.r.t. $\mathbf{\Gamma}$ the true ergodic capacity $\mathcal{C}_{mmse}(\mathbf{U}\mathbf{D}^{-1/2}\mathbf{\Gamma})$. This function of $\mathbf{\Gamma}$, as well as its derivative w.r.t. the entries of $\mathbf{\Gamma}$ cannot be expressed in closed form. They have to be evaluated by Monte Carlo simulations, thus complicating a lot the maximization algorithm.

APPENDIX

We prove that diagonal matrices \mathbf{R} are solutions of Problem 2. For this, we introduce a simplification of Problem 2 based on the observation that there exists a one-to-one correspondence between the set of matrices \mathbf{R} satisfying the constraints (7.12) and the set of positive matrices \mathbf{S} verifying

$$\mathbf{S} \geq 0, \quad \frac{\sigma^2}{p} \frac{1}{t} \text{Tr} \mathbf{D}^{-1} \mathbf{S} + \frac{1}{t} \text{Tr} \mathbf{S} (\mathbf{I} + \mathbf{S})^{-1} = \frac{r}{t} \quad (7.13)$$

In effect, if \mathbf{R} satisfies (7.12), then $\mathbf{S} = \delta(\mathbf{R})\mathbf{R}$ verifies

$$\frac{1}{t} \text{Tr} \mathbf{D}^{-1} \mathbf{S} = \delta(\mathbf{R}) \frac{1}{t} \text{Tr} \mathbf{D}^{-1} \mathbf{R} = \delta(\mathbf{R})$$

The equation (7.13) follows immediately from the definition 7.5 of $\delta(\mathbf{R})$. Conversely, if \mathbf{S} is a positive matrix for which (7.13) holds, then $\mathbf{R} = \frac{1}{\frac{1}{t} \text{Tr} \mathbf{D}^{-1} \mathbf{S}} \mathbf{S}$ satisfies (7.12) and $\delta(\mathbf{R}) = \frac{1}{t} \text{Tr} \mathbf{D}^{-1} \mathbf{S}$. Problem 2 is eventually equivalent to

Problem 3. *Maximize $\log \det (\mathbf{I} + \mathbf{S})$ under the constraints (7.13).*

and we have to establish that diagonal matrices are solutions of Problem 3. Let \mathbf{S} be a matrix verifying the constraints (7.13), and let $\mathbf{S} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^H$ be the eigenvalue/ eigenvector decomposition of matrix \mathbf{S} , where the eigenvalues $(\lambda_j)_{j=1,\dots,t}$ (the diagonal entries of diagonal matrix $\mathbf{\Lambda}$) are arranged in decreasing order. Then, we claim that it exists $\mu_d \geq 1$ such that $\mu_d \mathbf{\Lambda}$ verifies the constraints (7.13), and

$$\log_2 \det (\mathbf{I} + \mathbf{S}) = \log_2 \det (\mathbf{I} + \mathbf{\Lambda}) \leq \log_2 \det (\mathbf{I} + \mu_d \mathbf{\Lambda}) \quad (7.14)$$

To prove this, we first note that $\frac{1}{t}\text{Tr}\mathbf{S}(\mathbf{I} + \mathbf{S})^{-1} = \frac{1}{t}\text{Tr}\mathbf{\Lambda}(\mathbf{I} + \mathbf{\Lambda})^{-1}$. We recall that the eigenvalues $(d_j)_{j=1,\dots,t}$ are arranged in decreasing order. We can therefore use the useful inequality

$$\frac{1}{t}\text{Tr}\mathbf{D}^{-1}\mathbf{V}\mathbf{\Lambda}\mathbf{V}^H \geq \frac{1}{t}\text{Tr}\mathbf{D}^{-1}\mathbf{\Lambda} \quad (7.15)$$

which is extensively used in [32] (see the Appendix of [32]). (8.47) implies that

$$\frac{\sigma^2}{p} \frac{1}{t} \text{Tr}\mathbf{D}^{-1}\mathbf{\Lambda} + \frac{1}{t} \text{Tr}\mathbf{\Lambda}(\mathbf{I} + \mathbf{\Lambda})^{-1} \leq \frac{r}{t}$$

The function $\phi(\mu)$ defined by

$$\phi(\mu) = \frac{\sigma^2}{p} \frac{1}{t} \text{Tr}\mathbf{D}^{-1}\mu\mathbf{\Lambda} + \frac{1}{t} \text{Tr}\mu\mathbf{\Lambda}(\mathbf{I} + \mu\mathbf{\Lambda})^{-1}$$

is an increasing function such that $\phi(1) \leq \frac{r}{t}$. Therefore, it exists a unique $\mu_d \geq 1$ for which $\phi(\mu_d) = 1$. This means that matrix $\mu_d\mathbf{\Lambda}$ satisfies (7.13). Finally, equality (7.14) holds because $\mu_d \geq 1$. We denote by \mathbf{S}_d the matrix $\mathbf{S}_d = \mu_d\mathbf{\Lambda}$. (7.14) implies that for each matrix \mathbf{S} satisfying (7.13), then the diagonal matrix \mathbf{S}_d verifies (7.13) and $\log_2 \det(\mathbf{I} + \mathbf{S}) \leq \log_2 \det(\mathbf{I} + \mathbf{S}_d)$. This, in turn, establishes that diagonal matrices are solutions of Problem 3, and therefore of Problem 2.

$s \leq t$ components of a solution $\boldsymbol{\gamma}_{opt} = (\gamma_{1,opt}, \dots, \gamma_{t,opt})$ of Problem 1 all coincide with $\frac{t}{s}$, and the other $t - s$ are zero. The actual value of s depends on the signal to noise ratio. We note that if $s < t$, then,

Due to the lack of space, we do not develop this point in this paper.

Except in this simple situation, it seems difficult to characterize analytically the solutions

It is therefore useful to study Problem 1 both on the theoretical and on the practical point of view. In order to proceed to the theoretical study of Problem 1, we first slightly modify (as in the Appendix) the formulation of Problem 1. More precisely, it exists a one to one correspondance between the set of diagonal matrices $\boldsymbol{\Gamma}$ satisfying the constraints (8.44) and the set of diagonal matrices $\mathbf{\Lambda}$ verifying

$$\mathbf{\Lambda} \geq 0, \quad \frac{\sigma^2}{p} \frac{1}{t} \text{Tr}\mathbf{D}^{-1}\mathbf{\Lambda} + \frac{1}{t} \text{Tr}\mathbf{\Lambda}(\mathbf{I} + \mathbf{\Lambda})^{-1} = \frac{r}{t} \quad (7.16)$$

In effect, if $\boldsymbol{\Gamma}$ satisfies (8.44), then $\mathbf{\Lambda} = \delta(\boldsymbol{\Gamma})\boldsymbol{\Gamma}$ verifies

$$\frac{1}{t} \text{Tr}\mathbf{D}^{-1}\mathbf{\Lambda} = \delta(\boldsymbol{\Gamma}) \frac{1}{t} \text{Tr}\mathbf{D}^{-1}\boldsymbol{\Gamma} = \delta(\boldsymbol{\Gamma})$$

The equation (7.16) follows immediately from the definition (7.5) of $\delta(\boldsymbol{\Gamma})$. Conversely, if $\mathbf{\Lambda}$ is a positive diagonal matrix for which (7.16) holds, then $\boldsymbol{\Gamma} = \frac{1}{\frac{1}{t}\text{Tr}\mathbf{D}^{-1}\mathbf{\Lambda}} \mathbf{\Lambda}$ satisfies

(8.44) and $\delta(\mathbf{\Gamma}) = \frac{1}{t} \text{Tr } \mathbf{D}^{-1} \mathbf{\Lambda}$. The theoretical study of Problem 1 is thus equivalent to the study of Problem 4 defined by

Problem 4. *Maximize $\sum_{j=1}^t \log_2(1 + \lambda_j)$ under the constraints (7.16).*

Let $\boldsymbol{\lambda}_{opt} = (\lambda_{1,opt}, \dots, \lambda_{t,opt})$ be a solution of (4), and $\mathbf{\Lambda}_{opt}$ the associated positive diagonal matrix. We remark that $\boldsymbol{\lambda}_{opt}$ has to satisfy the Karush-Kuhn-Tucker conditions. In other words, it exists a Lagrange multiplier μ_{opt} such that

$$\begin{aligned} \frac{\partial \mathcal{L}(\mu_{opt})}{\partial \lambda_{l,opt}} &= 0 \text{ if } \lambda_{l,opt} > 0 \\ \frac{\partial \mathcal{L}(\mu_{opt})}{\partial \lambda_{l,opt}} &\leq 0 \text{ if } \lambda_{l,opt} = 0 \end{aligned} \quad (7.17)$$

where $\mathcal{L}(\mu_{opt})$ is defined by

$$\mathcal{L}(\mu_{opt}) = \sum_{j=1}^t \log_2(1 + \lambda_{j,opt}) - \frac{1}{\mu_{opt}} \left(\frac{\sigma^2}{p} \text{Tr } \mathbf{D}^{-1} \mathbf{\Lambda}_{opt} + \text{Tr } \mathbf{\Lambda}_{opt} (\mathbf{I} + \mathbf{\Lambda}_{opt})^{-1} - r \right)$$

It is difficult to evaluate $\lambda_{j,opt}$ in general

Chapitre 8

Article soumis à IEEE Information Theory

Sommaire

8.1	Introduction	113
8.2	Problem statement.	117
8.3	Derivation of the large system approximation of I_{mmse} . .	118
8.4	Structure of optimal precoders.	125
8.5	Maximization of \bar{I}_{mmse}	131
8.6	Concluding remarks.	138
8.7	Proof of Proposition 3.	138
8.8	Proof of Lemma 1.	149

On the precoder design of flat fading MIMO systems equipped with MMSE receivers : a large system approach.

Cedric Artigue, Philippe Loubaton

8.1 Introduction

It is now well established that using multiple transmit and receive antennas potentially allows to increase the Shannon capacity of digital communications systems. Since the seminal work of Teletar ([1]), the ergodic Shannon capacity of block fading MIMO systems has been studied extensively and important questions related to the design of optimal precoding schemes have been addressed. Considering that the Channel State Information (CSI) is available at receiver side while the transmitter is only aware of its second order statistical properties, many authors have studied the impact of antenna correlation on the capacity of MIMO systems communicating through flat fading channel ([32], [33]) and frequency selective channel ([34]).

The ergodic Shannon capacity is certainly a valuable figure of merit if the MIMO system under consideration is equipped with a maximum likelihood decoder. As the practical implementation of this decoder requires a high computational cost, it is also useful to study potential performance of MIMO systems equipped with the MMSE receiver. The corresponding (Gaussian) ergodic mutual information (EMI), denoted I_{mmse} in the following, is defined as the sum over the transmit antennas of the terms $\mathbb{E}(\log(1 + \beta_j))$, where β_j represents the output MMSE SINR associated to the stream sent by antenna j . The design of precoders maximizing I_{mmse} is of course an important issue because the optimum value of I_{mmse} represents the maximum rate that can be transmitted reliably when the MIMO system uses the MMSE receiver. This optimization problem has been extensively studied in the past, mainly if the CSI is available at the both the receiver and the transmitter (see e.g. [8]). It is however

often unrealistic to assume the CSI available at the transmitter side in the context of mobile systems.

In the present paper, we consider a flat fading MIMO channel with separable correlation structure (Kronecker model). We assume that the channel matrix is known at the receiver side, but that only its transmit and receive covariance matrices are available at the transmitter side. We address the problem of designing precoders that maximize I_{mmse} . The expression of I_{mmse} is rather complicated and thus difficult to maximize w.r.t. the precoding matrix. In particular, it seems difficult to establish that the left eigenvectors of an optimal precoding matrix coincide with the eigenvectors of the transmit correlation matrix as in the context of the evaluation of the Shannon ergodic capacity (see e.g. ([32])). Therefore, it is necessary to evaluate numerically both the singular values and the singular vectors of optimum precoding matrices, or equivalently to solve a t^2 dimensional optimization problem. Steepest descent algorithms require the use intensive Monte Carlo simulation technics in order to evaluate the gradient and/or the Hessian of the cost function (see e.g. [35] in the context of the evaluation of the Shannon capacity of correlated Rician channels). Moreover, the convergence of these algorithms is not guaranteed because I_{mmse} is in general not concave. As in previous contributions addressing the behaviour of the Shannon capacity of MIMO systems ([10], [11], [12]), we propose to replace the maximization of I_{mmse} by the maximization of an approximation obtained in the asymptotic regime $t \rightarrow \infty$, $r \rightarrow \infty$, $\frac{t}{r} \rightarrow c$, $c \in (0, \infty)$.

Large system approximation of I_{mmse} was previously considered in the context of CDMA systems with i.i.d. spreading codes (see e.g. [36] and the references herein), which, in the downlink, are formally equivalent to a subclass of the MIMO systems considered in this paper when the spreading codes are Gaussian. The specific case of MIMO systems has also been considered (see e.g. [37], [10]). It was shown that the SINRs $(\beta_j)_{j=1,\dots,t}$ converge towards deterministic terms depending on the transmit and receive covariance matrices (or their equivalent in the context of downlink CDMA systems). These results provide an obvious large system approximation \hat{I}_{mmse} of I_{mmse} .

In this paper, we establish that the large system approximation \hat{I}_{mmse} provides

a $O(\frac{1}{t})$ relative error. This is a rather poor convergence rate compared to the large system approximations of the Shannon capacity whose relative errors are $O(\frac{1}{t^2})$ ([15], [16]). We therefore propose to use an improved large system approximation, denoted \bar{I}_{mmse} , first introduced in [17] in the case of independent identically distributed (i.i.d.) MIMO channels, and then generalized independently in the conference papers [38] and [18]. The derivations of [18] are based on the replica method, a useful and powerful trick whose mathematical relevance has not yet been established in the present context, and thus differ from the large random matrix approach sketched in [38]. We show that the relative error associated to \bar{I}_{mmse} is a $O(\frac{1}{t^2})$ term, thus improving the predictions of [38] ($O(\frac{1}{t^{3/2}})$) and [18] ($o(\frac{1}{t})$). The method we use to study the accuracy of \bar{I}_{mmse} differs from [17] whose approach is somewhat similar to [39], a paper devoted to the asymptotic study of the SINRs $(\beta_j)_{j=1,\dots,t}$. The transmit covariance matrices of the MIMO channels of [39] are diagonal. This assumption simplifies the analysis so that the approach of [17], [39] cannot be generalized to the case of general transmit covariance matrices. Next, we address the maximization of \bar{I}_{mmse} w.r.t. the precoding matrix. We establish that the left singular vectors of an optimum precoder are the eigenvectors of the transmit covariance matrix and that its right eigenvectors matrix is equal to \mathbf{I}_t . The evaluation of a precoding matrix thus reduces to the evaluation of its singular values, a t -dimensional optimization problem. In general, the optimum singular values have no closed form expression. In order to get more insights on the optimum precoders, we consider the case of an uncorrelated MIMO channel for which it is possible to obtain in closed form the precoders which optimize the approximation \hat{I}_{mmse} . We show that the optimum precoders are the diagonal matrices whose entries are either 0, either all coincide with $\frac{t}{s}$ where s is the number of non zero entries which depend on the signal to noise ratio. Therefore, the optimum transmission strategy coincides with an antenna selection scheme. Although it is not proved that the above strategy maximizes I_{mmse} , this result shows that, at least if t is large enough, antenna selection may provide higher mutual informations I_{mmse} than a uniform power allocation. The situation differs from what was shown initially by Telatar ([1]) in the context of the study of the Shannon ergodic capacity of i.i.d. channels : the Shannon capacity achieving covariance matrix coincides with \mathbf{I}_t . We also remark that our result establishes formally

1. The authors wish to thank Aris Moustakas for suggesting that the rate $O(\frac{1}{t^{3/2}})$ was probably pessimistic

that \hat{I}_{mmse} is in general not a concave function of the precoding matrix, and infer from this that I_{mmse} is not concave as well. We finally consider the case of an arbitrary bicornelated MIMO channel, and propose to evaluate the singular values of an optimum precoder using a classical gradient algorithm. Numerical results show that the precoding matrices evaluated by this algorithm provide nearly the same mutual informations as direct approaches maximizing I_{mmse} while being computationally more attractive.

This paper is organized as follows. Section 8.2 is devoted to presentation of the problem and to the underlying assumptions. In section 8.3, we present the large system approximations \hat{I}_{mmse} and \bar{I}_{mmse} of I_{mmse} and analyse their accuracies. Section 8.4 studies the structure of the optimum precoders, and Section 8.5 addresses the optimization of \bar{I}_{mmse} .

General Notations In this paper, the notations y , \mathbf{x} , \mathbf{M} stand for scalars, vectors and matrices, respectively. As usual, $\|\mathbf{x}\|$ represents the Euclidian norm of vector \mathbf{x} and $\|\mathbf{M}\|$ stands for the spectral norm of matrix \mathbf{M} . The superscripts $(.)^T$ and $(.)^H$ represent respectively the transpose and transpose conjugate. The trace of \mathbf{M} is denoted by $\text{Tr}(\mathbf{M})$. The mathematical expectation operator is denoted by $\mathbb{E}(\cdot)$. The symbols \Re and \Im denote respectively the real and imaginary parts of a given complex number. If x is a possibly complex-valued random variable, $\text{Var}(x) = \mathbb{E}|x|^2 - |\mathbb{E}(x)|^2$ represents the variance of x .

All along this paper, t and r stand for the number of transmit and receive antennas. Certain quantities will be studied in the asymptotic regime $t \rightarrow \infty, r \rightarrow \infty$ in such a way that $\frac{t}{r} \rightarrow c \in (0, \infty)$. In order to simplify the notations, $t \rightarrow \infty$ should be understood from now on as $t \rightarrow \infty, r \rightarrow \infty$ and $\frac{t}{r} \rightarrow c \in (0, \infty)$. A vector \mathbf{x}_t and a matrix \mathbf{M}_t whose size depend on t are said to be uniformly bounded if $\sup_t \|\mathbf{x}_t\| < \infty$ and $\sup_t \|\mathbf{M}_t\| < \infty$.

Several variables used throughout this paper depend on various parameters, e.g. the number of antennas, the noise level, etc. In order to simplify the notations, we do not mention all these dependencies.

Notation C will denote a generic strictly positive constant whose main feature is not to depend on t . The value of C might change from one line to another.

8.2 Problem statement.

We consider a MIMO system equipped with r receive antennas and t transmit antennas. The MIMO channel matrix \mathbf{H} is supposed to be a Gaussian random matrix defined by

$$\mathbf{H} = \frac{1}{\sqrt{t}} \mathbf{C}_R^{1/2} \mathbf{H}_{iid} \mathbf{C}_T^{1/2} \quad (8.1)$$

where \mathbf{H}_{iid} is a rt matrix whose entries are independent and identically distributed (i.i.d.) complex circular Gaussian random variables $\mathcal{CN}(0, 1)$, i.e. $H_{iid,ij} = \Re H_{iid,ij} + \mathbf{i} \Im H_{iid,ij}$ where $\Re H_{iid,ij}$ and $\Im H_{iid,ij}$ are independent centered real Gaussian random variables with variance $\frac{1}{2}$. Matrices \mathbf{C}_T and \mathbf{C}_R are positive definite matrices modeling respectively the impact of correlation between transmitting and receiving antennas. We assume that $\frac{1}{t} \text{Trace}(\mathbf{C}_T) = 1$ and $\frac{1}{r} \text{Trace}(\mathbf{C}_R) = 1$. This assumption implies that $\frac{1}{r} \mathbb{E}(\text{Tr}(\mathbf{H}\mathbf{H}^H)) = 1$.

Each transmit antenna j sends a sequence $(x_j(n))_{n \in \mathbb{Z}}$ defined by

$$\mathbf{x}(n) = (x_1(n), \dots, x_t(n))^T = \mathbf{K}\mathbf{s}(n) = \mathbf{K}(s_1(n), \dots, s_t(n))^T$$

where the $((s_j(n))_{n \in \mathbb{Z}})_{j=1, \dots, t}$ are assumed to be unit variance mutually independent i.i.d. sequences. \mathbf{K} represents a precoding matrix satisfying $\frac{1}{t} \text{Tr}(\mathbf{K}\mathbf{K}^H) \leq 1$.

The corresponding r -variate discrete-time received signal $(\mathbf{y}(n))_{n \in \mathbb{Z}}$ is given by

$$\mathbf{y}(n) = \mathbf{H}\mathbf{K}\mathbf{s}(n) + \mathbf{n}(n) \quad (8.2)$$

where \mathbf{n} is a white Gaussian noise with covariance matrix $\mathbf{E}(\mathbf{n}(n)\mathbf{n}(n)^H) = \sigma^2 \mathbf{I}_r$.

In this paper, we evaluate the potential performance of the MIMO system (8.2) when the receiver is equipped with the MMSE receiver. In other words, each symbol sequence s_j is estimated by the Wiener filter prior to decoding, i.e. $s_j(n)$ is estimated by

$$\hat{s}_j(n) = \mathbf{k}_j^H \mathbf{H}^H (\mathbf{H}\mathbf{K}\mathbf{K}^H \mathbf{H}^H + \sigma^2 \mathbf{I}_r)^{-1} \mathbf{y}(n)$$

where \mathbf{k}_j represents the column j of \mathbf{K} . In the following, we denote by $\mathbf{Q}_T(\mathbf{K})$ the matrix

$$\mathbf{Q}_T(\mathbf{K}) = (\mathbf{K}^H \mathbf{H}^H \mathbf{H} \mathbf{K} + \sigma^2 \mathbf{I}_t)^{-1} \quad (8.3)$$

It is standard that the SINR β_j provided by this linear receiver is given by ([40])

$$\beta_j(\mathbf{K}) = \frac{1}{\sigma^2 (\mathbf{Q}_T(\mathbf{K}))_{j,j}} - 1 \quad (8.4)$$

The ergodic mutual information $I_{mmse}(\mathbf{K})$ of the MIMO system under consideration is thus equal to

$$I_{mmse}(\mathbf{K}) = \mathbb{E} \left[\sum_{j=1}^t \log(1 + \beta_j(\mathbf{K})) \right] = -\mathbb{E} \left[\sum_{j=1}^t \log(\sigma^2 \mathbf{Q}_T(\mathbf{K}))_{j,j} \right] \quad (8.5)$$

where the mathematical expectation is over the probability distribution of random matrix \mathbf{H} . In order to maximize $I_{mmse}(\mathbf{K})$ over the set $\frac{1}{t} \text{Tr}(\mathbf{K}\mathbf{K}^H) \leq 1$, it is necessary to use numerical technics based on steepest descent algorithms. As the gradient and the Hessian of I_{mmse} have no simple expression, they have to be evaluated using intensive Monte Carlo simulations (see e.g. [35]). Moreover, to our best knowledge, the singular vectors of an optimum matrix have no closed form expression. Therefore, the dimension of the optimization problem cannot be reduced from t^2 to t as in the context of the evaluation of the capacity achieving covariance matrix ([32]).

8.3 Derivation of the large system approximation of I_{mmse} .

In this section, we introduce the large system approximation presented in [38] and [18], and improve the results stated without proof in [38] concerning its accuracy. Our approach is based on Gaussian large random matrix technics initiated by Pastur ([41]). Pastur's approach was used in [16] in order to establish the asymptotic Gaussianity of the traditional mutual information of bicorrelated MIMO channels.

We study in this section the asymptotic behaviour of I_{mmse} in the case where the precoding matrix \mathbf{K} is reduced to $\mathbf{K} = \mathbf{I}_t$ to simplify the notations. In order to deduce the results in the case $\mathbf{K} \neq \mathbf{I}_t$, we remark that channel matrix $\mathbf{H}\mathbf{K}$ can be interpreted as a bi-correlated MIMO channel with transmit and receive covariance matrices $\mathbf{K}^H \mathbf{C}_T \mathbf{K}$ and \mathbf{C}_R respectively. We will therefore replace matrix \mathbf{C}_T by matrix $\mathbf{K}^H \mathbf{C}_T \mathbf{K}$. $I_{mmse}(\mathbf{I})$ and $\mathbf{Q}_T(\mathbf{I})$ are denoted I_{mmse} and \mathbf{Q}_T in the remainder of this section.

We first explain the differences between our analysis and the contributions [17] and [39]. We recall that [17] addresses the i.i.d. case while [39] assumes that matrix $\mathbf{C}_T = \text{Diag}(c_{T,1}, \dots, c_{T,t})$ is diagonal. In this last context, the SINR β_j can also be

written as

$$\beta_j = \frac{c_{T,j}}{t} \mathbf{h}_{iid,j}^H \left(\mathbf{C}_R^{1/2} \mathbf{H}_{iid}^{(j)} \frac{\mathbf{C}_T^{(j)}}{t} \mathbf{H}_{iid}^{(j)H} \mathbf{C}_R^{1/2} \right)^{-1} \mathbf{h}_{iid,j} = \frac{c_{T,j}}{t} \mathbf{h}_{iid,j}^H \mathbf{Q}_R^{(j)} \mathbf{h}_{iid,j} \quad (8.6)$$

where $\mathbf{h}_{iid}^{(j)}$ represents the column j of \mathbf{H}_{iid} , matrix $\mathbf{H}_{iid}^{(j)}$ is obtained from \mathbf{H}_{iid} by deleting column j , and $\mathbf{C}_T^{(j)}$ represents the $(t-1) \times (t-1)$ diagonal matrix obtained by deleting column and row j from \mathbf{C}_T . The approaches of [17] and [39] rely on the key observation that vector $\mathbf{h}_{iid,j}$ is independent from the matrix $\mathbf{Q}_R^{(j)}$. This allows to study the behaviour of β_j using important results concerning the behaviour of random quadratic forms. If matrix \mathbf{C}_T is non diagonal, β_j has not the same structure than in (8.6) : vector $\mathbf{h}_{iid}^{(j)}$ and matrices $\mathbf{Q}_R^{(j)}$ are replaced by non independent terms, and the approach of [17] and [39] cannot be used. Our approach does not study β_j directly, but rather the diagonal entries of matrix $\sigma^2 \mathbf{Q}_T$ whose asymptotic behaviour can be evaluated for general transmit covariance matrices \mathbf{C}_T .

The study of the accuracy of the approximation is essentially based on the study of a virtual channel obtained from \mathbf{H} after unitary transformations. We consider the eigenvalue/ eigenvector decompositions of covariance matrices \mathbf{C}_T and \mathbf{C}_R :

$$\mathbf{C}_T = \mathbf{U} \mathbf{D} \mathbf{U}^H, \quad \mathbf{C}_R = \tilde{\mathbf{U}} \tilde{\mathbf{D}} \tilde{\mathbf{U}}^H \quad (8.7)$$

where the diagonal entries $(d_j)_{j=1,\dots,t}$ and $(\tilde{d}_i)_{i=1,\dots,r}$ of \mathbf{D} and $\tilde{\mathbf{D}}$ are arranged in the decreasing order. Then, we define the random $t \times r$ matrix \mathbf{Y} by

$$\mathbf{Y}^H = \tilde{\mathbf{U}}^H \mathbf{H} \mathbf{U} \quad (8.8)$$

\mathbf{Y} can be written as

$$\mathbf{Y} = \frac{1}{\sqrt{t}} \mathbf{D}^{1/2} \mathbf{X} \tilde{\mathbf{D}}^{1/2} \quad (8.9)$$

where \mathbf{X} represents the $t \times r$ matrix $\mathbf{X} = \mathbf{U}^H \mathbf{H}_{iid} \tilde{\mathbf{U}}$. As \mathbf{U} and $\tilde{\mathbf{U}}$ are unitary, matrix \mathbf{X} is an i.i.d. complex Gaussian matrix such that $\mathbb{E}|X_{i,j}|^2 = 1$. In the following, we denote by \mathbf{Q} the matrix defined by

$$\mathbf{Q} = (\mathbf{Y} \mathbf{Y}^H + \sigma^2 \mathbf{I})^{-1} \quad (8.10)$$

The study of I_{mmse} when $t \rightarrow \infty$ is based on the asymptotic properties of the diagonal entries of matrix \mathbf{Q}_T . We remark that

$$\mathbf{Q}_T = \mathbf{U} \mathbf{Q} \mathbf{U}^H \quad (8.11)$$

and evaluate the asymptotic behaviour of $\mathbf{u}\mathbf{Q}\mathbf{u}^H$ where $\mathbf{u} = (u_1, \dots, u_t)$ is a unit norm deterministic row vector. We use in the following certain results of [16]. We however note that in [16], matrix \mathbf{Q} is replaced by matrix $(\mathbf{I} + \sigma^2 \mathbf{Y}\mathbf{Y}^H)^{-1}$. Therefore, the statements of [16] have to be adapted. In the sequel, we denote by δ and $\tilde{\delta}$ the unique strictly positive solutions of the system

$$\begin{aligned}\delta &= \frac{1}{t} \text{Tr} \left[\mathbf{D} \left(\sigma^2 (\mathbf{I} + \tilde{\delta} \mathbf{D}) \right) \right]^{-1} \\ \tilde{\delta} &= \frac{1}{t} \text{Tr} \left[\tilde{\mathbf{D}} \left(\sigma^2 (\mathbf{I} + \delta \tilde{\mathbf{D}}) \right) \right]^{-1}\end{aligned}\tag{8.12}$$

The existence and the uniqueness of the solution has been established in Proposition 1 of [16]. We denote by \mathbf{T} and $\tilde{\mathbf{T}}$ the diagonal matrices

$$\begin{aligned}\mathbf{T} &= \left[\sigma^2 (\mathbf{I} + \tilde{\delta} \mathbf{D}) \right]^{-1} \\ \tilde{\mathbf{T}} &= \left[\sigma^2 (\mathbf{I} + \delta \tilde{\mathbf{D}}) \right]^{-1}\end{aligned}\tag{8.13}$$

and gather in the following proposition certain useful results of [16].

Proposition 2. *Assume that matrices \mathbf{D} and $\tilde{\mathbf{D}}$ satisfy the following conditions :*

$$\begin{aligned}\sup_t \|\mathbf{D}\| &\leq d_{\max} < \infty \quad , \quad \inf_t \frac{1}{t} \text{Tr} \mathbf{D} > 0 \\ \sup_t \|\tilde{\mathbf{D}}\| &\leq \tilde{d}_{\max} < \infty \quad , \quad \inf_t \frac{1}{t} \text{Tr} \tilde{\mathbf{D}} > 0\end{aligned}\tag{8.14}$$

Then, the following results hold true :

- For each uniformly bounded deterministic matrix \mathbf{M}^2

$$\begin{aligned}\text{Var} \left(\frac{1}{t} \text{Tr} (\mathbf{M}\mathbf{Q}) \right) &= O\left(\frac{1}{t^2}\right) \\ \mathbb{E} \left(\text{Tr} (\mathbf{M}(\mathbf{Q} - \mathbf{T})) \right) &= O\left(\frac{1}{t}\right)\end{aligned}\tag{8.15}$$

- γ and $\tilde{\gamma}$ defined by

$$\gamma = \frac{1}{t} \text{Tr} (\mathbf{D}^2 \mathbf{T}^2) \quad , \quad \tilde{\gamma} = \frac{1}{t} \text{Tr} (\tilde{\mathbf{D}}^2 \tilde{\mathbf{T}}^2) \quad ,\tag{8.16}$$

satisfy

$$\inf_t (1 - \sigma^4 \gamma \tilde{\gamma}) > 0\tag{8.17}$$

We assume from now on that the matrices \mathbf{D} and $\tilde{\mathbf{D}}$ satisfy (8.14). We are now in position to state the main results of this section. We begin by the following proposition.

2. In [16], matrix \mathbf{M} is diagonal. The case of non diagonal matrices is addressed in [42] devoted to correlated Ricean channels.

Proposition 3.

$$\sup_{\mathbf{u}, \|\mathbf{u}\|=1} |\mathbb{E}(\mathbf{u}(\mathbf{Q} - \mathbf{T})\mathbf{u}^H)| = O\left(\frac{1}{t^{3/2}}\right) \quad (8.18)$$

and

$$\sup_{\mathbf{u}, \|\mathbf{u}\|=1} \left| \mathbb{E}(\mathbf{u}(\mathbf{Q} - \mathbb{E}(\mathbf{Q}))\mathbf{u}^H)^3 \right| = O\left(\frac{1}{t^2}\right) \quad (8.19)$$

Moreover,

$$\sup_{\mathbf{u}, \|\mathbf{u}\|=1} \left| \text{Var}(\mathbf{u}\mathbf{Q}\mathbf{u}^H) - \frac{1}{t} \frac{\sigma^4 \tilde{\gamma}}{1 - \sigma^4 \gamma \tilde{\gamma}} (\mathbf{u}\mathbf{T}^2 \mathbf{D} \mathbf{u}^H)^2 \right| = O\left(\frac{1}{t^{3/2}}\right) \quad (8.20)$$

Finally, if we denote by $(\mathbf{v}_k)_{k=1, \dots, t}$ the row vectors of any unitary matrix \mathbf{V} , and if $(\kappa_j)_{j=1, \dots, t}$ denote positive numbers such that $\sup_j \kappa_j < C$, we have

$$\sum_{k=1}^t \left[\text{Var}(\kappa_k \mathbf{v}_k \mathbf{Q} \mathbf{v}_k^H) - \frac{1}{t} \frac{\sigma^4 \tilde{\gamma}}{1 - \sigma^4 \gamma \tilde{\gamma}} (\kappa_k \mathbf{v}_k \mathbf{T}^2 \mathbf{D} \mathbf{v}_k^H)^2 \right] = O\left(\frac{1}{t}\right) \quad (8.21)$$

The proof is given in the Appendix. In order to introduce the large system approximation \bar{I}_{mmse} , we define matrices \mathbf{T}_T and \mathbf{T}_R by

$$\mathbf{T}_T = \mathbf{U} \mathbf{T} \mathbf{U}^H = \left(\sigma^2 (\mathbf{I} + \tilde{\delta} \mathbf{C}_T) \right)^{-1}, \quad \mathbf{T}_R = \tilde{\mathbf{U}} \tilde{\mathbf{T}} \tilde{\mathbf{U}}^H = \left(\sigma^2 (\mathbf{I} + \delta \mathbf{C}_R) \right)^{-1} \quad (8.22)$$

We note that $(\delta, \tilde{\delta})$ and $(\gamma, \tilde{\gamma})$ can be expressed in terms of \mathbf{T}_T and \mathbf{T}_R as

$$\begin{aligned} \delta &= \frac{1}{t} \text{Tr} \left[\mathbf{C}_T \left(\sigma^2 (\mathbf{I} + \tilde{\delta} \mathbf{C}_T) \right)^{-1} \right] = \frac{1}{t} \text{Tr} [\mathbf{C}_T \mathbf{T}_T] \\ \tilde{\delta} &= \frac{1}{t} \text{Tr} \left[\mathbf{C}_R \left(\sigma^2 (\mathbf{I} + \delta \mathbf{C}_R) \right)^{-1} \right] = \frac{1}{t} \text{Tr} [\mathbf{C}_R \mathbf{T}_R] \end{aligned} \quad (8.23)$$

and

$$\gamma = \frac{1}{t} \text{Tr}(\mathbf{C}_T^2 \mathbf{T}_T^2), \quad \tilde{\gamma} = \frac{1}{t} \text{Tr}(\mathbf{C}_R^2 \mathbf{T}_R^2) \quad (8.24)$$

The following result holds.

Theorem 3. We define \bar{I}_{mmse} by

$$\bar{I}_{mmse} = - \sum_{j=1}^t \log(\sigma^2 \mathbf{T}_{T,j,j}) + \frac{1}{2\tilde{\delta}^2} \frac{\tilde{\gamma}}{1 - \sigma^4 \gamma \tilde{\gamma}} \frac{1}{t} \sum_{j=1}^t \left(1 - \frac{((\sigma^2 \mathbf{T}_T)^2)_{j,j}}{\sigma^2 \mathbf{T}_{T,j,j}} \right)^2 \quad (8.25)$$

Then,

$$I_{mmse} = \bar{I}_{mmse} + O\left(\frac{1}{t}\right) \quad (8.26)$$

Proof. The proof is based on a second order expansion of $\log \sigma^2 \mathbf{Q}_{T,j,j}$ around the point $\mathbb{E}(\mathbf{Q}_{T,j,j})$. We define ϵ_j by

$$\epsilon_j = \frac{\mathbf{Q}_{T,j,j} - \mathbb{E}(\mathbf{Q}_{T,j,j})}{\mathbb{E}(\mathbf{Q}_{T,j,j})}$$

and write $\log \sigma^2 \mathbf{Q}_{T,j,j}$ as

$$\log \sigma^2 \mathbf{Q}_{T,j,j} = \log (\sigma^2 \mathbb{E}(\mathbf{Q}_{T,j,j})) + \log(1 + \epsilon_j)$$

We express $\log(1 + \epsilon_j)$ as

$$\log(1 + \epsilon_j) = \epsilon_j - \frac{\epsilon_j^2}{2} + \frac{\epsilon_j^3}{3} + r_j$$

As $\mathbb{E}(\epsilon_j) = 0$, $\mathbb{E}(\log \sigma^2 \mathbf{Q}_{T,j,j})$ can be written as

$$\mathbb{E}(\log \sigma^2 \mathbf{Q}_{T,j,j}) = \log (\sigma^2 \mathbb{E}(\mathbf{Q}_{T,j,j})) - \frac{1}{2} \mathbb{E}(\epsilon_j^2) + \frac{1}{3} \mathbb{E}(\epsilon_j^3) + \mathbb{E}(r_j) \quad (8.27)$$

In order to be able to use Proposition 3, we have to study the behaviour $\frac{1}{\mathbb{E}(\mathbf{Q}_{T,j,j})}$. We first remark that it exists a deterministic constant $C > 0$ such that

$$\sup_{j,t} \frac{1}{\mathbb{E}(\mathbf{Q}_{T,j,j})} \leq C \quad (8.28)$$

Indeed, $\frac{1}{\mathbb{E}(\mathbf{Q}_{T,j,j})} \leq \mathbb{E}\left(\frac{1}{\mathbf{Q}_{T,j,j}}\right) = \sigma^2 \mathbb{E}(1 + \beta_j)$ by the Jensen inequality. We denote by \mathbf{h}_j the column j of matrix \mathbf{H} . The SINR β_j provided by the MMSE receiver is upperbounded by the match filter bound, i.e. $\beta_j \leq \frac{\|\mathbf{h}_j\|^2}{\sigma^2}$. As it is clear that $\sup_{j,t} \mathbb{E}(\|\mathbf{h}_j\|^2) \leq C$, we get (8.28).

We use (8.21) with $\kappa_j = \frac{1}{\mathbb{E}(\mathbf{Q}_{T,j,j})}$ and when the unitary matrix \mathbf{V} coincides with matrix \mathbf{U} . We obtain immediately that

$$\sum_{j=1}^t \mathbb{E}(\epsilon_j^2) - \frac{1}{t} \frac{\sigma^4 \tilde{\gamma}}{1 - \sigma^4 \gamma \tilde{\gamma}} \left(\frac{\mathbf{u}_j \mathbf{T}^2 \mathbf{D} \mathbf{u}_j^H}{\mathbb{E}(\mathbf{u}_j \mathbf{Q} \mathbf{u}_j^H)} \right)^2 = O\left(\frac{1}{t}\right) \quad (8.29)$$

We now establish that

$$\sup_j \left| \frac{1}{\mathbb{E}(\mathbf{Q}_{T,j,j})} - \frac{1}{\mathbf{T}_{T,j,j}} \right| \leq \frac{C}{t^{3/2}} \quad (8.30)$$

For this, we first notice that

$$\frac{1}{\mathbf{T}_{T,j,j}} \leq C \quad (8.31)$$

Indeed,

$$\frac{1}{\mathbf{T}_{T,j,j}} \leq (\mathbf{T}_T)_{j,j} = \sigma^2 \left(1 + \tilde{\delta} \mathbf{C}_{T,j,j}\right)$$

The conclusion follows $\tilde{\delta} \leq \frac{1}{\sigma^2} \frac{1}{t} \text{Tr}(\tilde{\mathbf{D}}) \leq \frac{r \tilde{d}_{max}}{t \sigma^2}$. (8.30) follows directly from

$$\frac{1}{\mathbb{E}(\mathbf{Q}_{T,j,j})} - \frac{1}{\mathbf{T}_{T,j,j}} = \frac{\mathbf{T}_{T,j,j} - \mathbb{E}(\mathbf{Q}_{T,j,j})}{\mathbf{T}_{T,j,j} \mathbb{E}(\mathbf{Q}_{T,j,j})}$$

and from (8.18).

(8.29) and (8.30) imply that

$$\sum_{j=1}^t \mathbb{E}(\epsilon_j^2) - \frac{1}{t} \frac{\sigma^4 \tilde{\gamma}}{1 - \sigma^4 \gamma \tilde{\gamma}} \left(\frac{\mathbf{u}_j \mathbf{T}^2 \mathbf{D} \mathbf{u}_j^H}{\mathbf{u}_j \mathbf{T} \mathbf{u}_j^H} \right)^2 = O\left(\frac{1}{t}\right) \quad (8.32)$$

Moreover, (8.19) and (8.28) lead to

$$\sup_j \mathbb{E}(\epsilon_j^3) \leq \frac{C}{t^2} \quad (8.33)$$

In order to evaluate the influence of r_j , we give the following lemma, proved in the appendix.

Lemma 1.

$$\sup_j |\mathbb{E}(r_j)| = O\left(\frac{1}{t^2}\right) \quad (8.34)$$

(8.32) and (8.34) imply that

$$I_{mmse} = - \sum_{j=1}^t \log \mathbb{E}(\sigma^2 \mathbf{Q}_{T,j,j}) + \frac{1}{2t} \frac{\sigma^4 \tilde{\gamma}}{1 - \sigma^4 \gamma \tilde{\gamma}} \sum_{j=1}^t \left(\frac{\mathbf{u}_j \mathbf{T}^2 \mathbf{D} \mathbf{u}_j^H}{\mathbf{u}_j \mathbf{T} \mathbf{u}_j^H} \right)^2 + O\left(\frac{1}{t}\right)$$

Straightforward manipulations show that

$$\sum_{j=1}^t \left(\frac{\mathbf{u}_j \mathbf{T}^2 \mathbf{D} \mathbf{u}_j^H}{\mathbf{u}_j \mathbf{T} \mathbf{u}_j^H} \right)^2 = \frac{1}{\tilde{\delta}^2} \sum_{j=1}^t \left(1 - \frac{((\sigma^2 \mathbf{T}_T)^2)_{j,j}}{\sigma^2 \mathbf{T}_{T,j,j}} \right)^2$$

In order to establish Theorem 3, it remains to prove that

$$\sum_{j=1}^t \log \mathbb{E}(\sigma^2 \mathbf{Q}_{T,j,j}) = \sum_{j=1}^t \log (\sigma^2 \mathbf{T}_{T,j,j}) + O\left(\frac{1}{t}\right) \quad (8.35)$$

We define $\bar{\epsilon}_j$ as

$$\bar{\epsilon}_j = \frac{\mathbf{T}_{T,j,j} - \mathbb{E}(\mathbf{Q}_{T,j,j})}{\mathbb{E}(\mathbf{Q}_{T,j,j})}$$

and remark that

$$\log(\sigma^2 \mathbf{T}_{T,j,j}) = \log \mathbb{E}(\sigma^2 \mathbf{Q}_{T,j,j}) + \log(1 + \bar{\epsilon}_j)$$

Using (8.28), we obtain that

$$|\bar{\epsilon}_j| \leq C |\mathbf{T}_{T,j,j} - \mathbb{E}(\mathbf{Q}_{T,j,j})| \quad (8.36)$$

(8.18) implies that $\sup_j |\mathbf{T}_{T,j,j} - \mathbb{E}(\mathbf{Q}_{T,j,j})| = O(\frac{1}{t^{3/2}})$. By (8.36), we get that $\sup_j |\bar{\epsilon}_j| = O(\frac{1}{t^{3/2}})$. For t large enough, $|\bar{\epsilon}_j| < A < 1$ for each j . For these t , we can write $\log(1 + \bar{\epsilon}_j)$ as

$$\log(1 + \bar{\epsilon}_j) = \bar{\epsilon}_j + \bar{r}_j$$

where

$$\bar{r}_j = (\bar{\epsilon}_j)^2 \sum_{n=2}^{\infty} (-1)^{n-1} \frac{\bar{\epsilon}_j^{n-2}}{n}$$

By (8.18), it holds that $\sup_j \bar{\epsilon}_j^2 = O(\frac{1}{t^3})$. Therefore,

$$|\bar{r}_j| < (\bar{\epsilon}_j)^2 \sum_{n=2}^{\infty} \frac{A^{n-2}}{n} < \frac{C}{t^3}$$

Consequently,

$$\sum_{j=1}^t \log(1 + \bar{\epsilon}_j) = \sum_{j=1}^t \bar{\epsilon}_j + O(\frac{1}{t^2})$$

We finally remark that

$$\sum_{j=1}^t \bar{\epsilon}_j = \sum_{j=1}^t \kappa_j \mathbf{u}_j \mathbb{E}(\mathbf{Q} - \mathbf{T}) \mathbf{u}_j^H$$

where $\kappa_j = \frac{1}{\mathbb{E}(\mathbf{Q}_{T,j,j})}$. The second item of (8.15) can thus be used for matrix $\mathbf{M} = \sum_j \kappa_j \mathbf{u}_j \mathbf{u}_j^H$, thus showing that (8.35) holds. This completes the proof of (8.26).

We denote \hat{I}_{mmse} the term defined by

$$\hat{I}_{mmse} = - \sum_{j=1}^t \log(\sigma^2 \mathbf{T}_{T,j,j}) \quad (8.37)$$

\hat{I}_{mmse} corresponds to the obvious large system approximation of I_{mmse} obtained by replacing, for each j , $(1 + \beta_j)$ by its "deterministic equivalent" $(\sigma^2 \mathbf{T}_{T,j,j})^{-1}$. Theorem 3 shows that the relative error provided by \hat{I}_{mmse} is a $O(\frac{1}{t})$ term, while the relative

error of \bar{I}_{mmse} is a $O(\frac{1}{t^2})$ term.

We now present some simulation experiments which demonstrate the accuracy of the approximation \bar{I}_{mmse} for a realistic number of antennas. \hat{I}_{mmse} is also represented. The transmit antennas correlation matrix \mathbf{C}_T is generated according to the popular model proposed in [9], i.e.

$$\mathbf{C}_{T,k,l} = a e^{-i\pi(k-l)\cos\phi_T} e^{-\frac{1}{2}(\pi(k-l)\sin\phi_T\sigma_{\phi_T})^2} \quad (8.38)$$

where a is a constant chosen in such a way that $\frac{1}{t}\text{Tr}(\mathbf{C}_T) = 1$. ϕ_T and σ_{ϕ_T} can be interpreted as the mean angle of departure and the standard deviation of the angles of departure of a scatterer cluster respectively. We notice that if $\sigma_{\phi_T} \simeq 0$, then $\text{Rank}(\mathbf{C}_T) \simeq 1$. We refer the reader to [9] for more details.

The receive antennas correlation matrix is generated similarly with different parameters ϕ_R and σ_{ϕ_R} .

In Figure 1 we have represented I_{mmse} , \hat{I}_{mmse} , \bar{I}_{mmse} versus the SNR for $r = t = 4$. Here, the various parameters are equal to $\phi_T = \pi/4$, $\sigma_{\phi_T} = 0.5$, $\phi_R = \pi/12$, $\sigma_{\phi_R} = 0.5$. We observe that \hat{I}_{mmse} can be rather far from the true mutual information I_{mmse} evaluated by Monte-Carlo simulation over 1000 channel realizations. Figure 2 represents the relative error between I_{mmse} and \hat{I}_{mmse} , \bar{I}_{mmse} respectively in terms of the mean angle of departure variance $\sigma_{\phi_T}^2$ for SNR = 0 dB and SNR = 6 dB when $\phi_T = \pi/4$, $\phi_R = \pi/12$, $\sigma_{\phi_R} = 0.4$. Figures 1,2 show that approximation \bar{I}_{mmse} provides significantly better results than \hat{I}_{mmse} .

The expression (8.25) is a large system approximation of $I_{mmse}(\mathbf{I})$. If the precoding matrix \mathbf{K} is not equal to \mathbf{I} , the approximation of $I_{mmse}(\mathbf{K})$ is obtained by replacing matrix \mathbf{C}_T by matrix $\mathbf{K}^H \mathbf{C}_T \mathbf{K}$. In the following we denote by $\delta(\mathbf{K})$, $\tilde{\delta}(\mathbf{K})$, $\mathbf{T}_T(\mathbf{K})$, $\mathbf{T}_R(\mathbf{K})$, $\gamma(\mathbf{K})$, $\tilde{\gamma}(\mathbf{K})$, $\bar{I}_{mmse}(\mathbf{K})$, $\hat{I}_{mmse}(\mathbf{K})$ the values of parameters δ , $\tilde{\delta}$, \mathbf{T}_T , \mathbf{T}_R , γ , $\tilde{\gamma}$, \bar{I}_{mmse} , \hat{I}_{mmse} when \mathbf{C}_T is replaced by $\mathbf{K}^H \mathbf{C}_T \mathbf{K}$.

8.4 Structure of optimal precoders.

In this section, we study the problem of designing precoders maximizing function $\bar{I}_{mmse}(\mathbf{K})$ over the set \mathcal{K} defined by

$$\mathcal{K} = \{\mathbf{K}, \frac{1}{t}\text{Tr}(\mathbf{K}\mathbf{K}^H) \leq 1\} \quad (8.39)$$

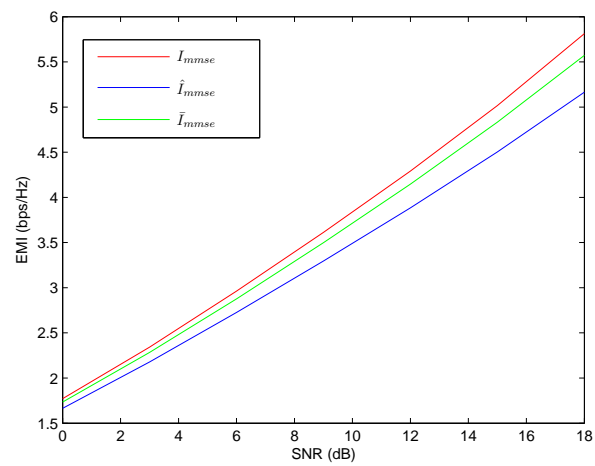


FIGURE 8.1 – Accuracy of the large system approximant

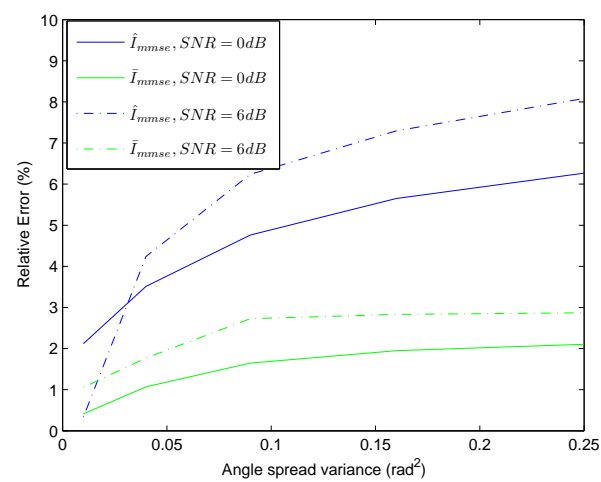


FIGURE 8.2 – Relative error

The main result of this section states that there is no restriction to look for optimal precoders of the form $\mathbf{K} = \mathbf{U}\mathbf{D}^{-1/2}\mathbf{\Lambda}^{1/2}$ where $\mathbf{\Lambda}$ is a diagonal matrix with positive elements. In order to establish this, we first derive the following intermediate result.

Proposition 4. *Let \mathbf{K} be an element of \mathcal{K} and the eigenvalue/eigenvector decomposition of matrix $\mathbf{K}^H\mathbf{C}_T\mathbf{K}$*

$$\mathbf{K}^H\mathbf{C}_T\mathbf{K} = \mathbf{W}\mathbf{\Lambda}\mathbf{W}^H$$

Then, matrix $\mathbf{K}_d = \mathbf{K}\mathbf{W}$ belongs to \mathcal{K} and satisfies

$$\bar{I}_{mmse}(\mathbf{K}) \leq \bar{I}_{mmse}(\mathbf{K}_d) \quad (8.40)$$

Proof. It is obvious that $\mathbf{K}_d \in \mathcal{K}$. In order to establish (8.40), we denote by $\bar{J}_{mmse}(\mathbf{K})$ the term

$$\bar{J}_{mmse}(\mathbf{K}) = \frac{1}{2\tilde{\delta}(\mathbf{K})^2} \frac{\tilde{\gamma}(\mathbf{K})}{1 - \sigma^4\gamma(\mathbf{K})\tilde{\gamma}(\mathbf{K})} \frac{1}{t} \sum_{j=1}^t \left(1 - \frac{(\sigma^2\mathbf{T}_T(\mathbf{K}))_{j,j}^2}{\sigma^2\mathbf{T}_{T,j,j}(\mathbf{K})} \right)^2 \quad (8.41)$$

and remark that $\bar{I}_{mmse}(\mathbf{K}) = \hat{I}_{mmse}(\mathbf{K}) + \bar{J}_{mmse}(\mathbf{K})$. We prove in the following that $\hat{I}_{mmse}(\mathbf{K}) \leq \hat{I}_{mmse}(\mathbf{K}_d)$ and $\bar{J}_{mmse}(\mathbf{K}) \leq \bar{J}_{mmse}(\mathbf{K}_d)$.

We first remark that $\mathbf{K}_d^H\mathbf{C}_T\mathbf{K}_d$ is the diagonal matrix $\mathbf{\Lambda}$. Therefore, by (8.22) matrix $\sigma^2\mathbf{T}_T(\mathbf{K}_d)$ is also diagonal, and is given by

$$\sigma^2\mathbf{T}_T(\mathbf{K}_d) = \left(\mathbf{I} + \tilde{\delta}(\mathbf{K}_d)\mathbf{\Lambda} \right)^{-1}$$

Moreover,

$$\sigma^2\mathbf{T}_T(\mathbf{K}) = \mathbf{W}\sigma^2\mathbf{T}_T(\mathbf{K}_d)\mathbf{W}^H \quad (8.42)$$

We claim that $(\delta(\mathbf{K}), \tilde{\delta}(\mathbf{K})) = (\delta(\mathbf{K}_d), \tilde{\delta}(\mathbf{K}_d))$. To check this, we recall that $(\delta(\mathbf{K}), \tilde{\delta}(\mathbf{K}))$ are defined as the unique positive solutions of the system

$$\begin{aligned} \delta(\mathbf{K}) &= \frac{1}{t} \text{Tr} \mathbf{K}^H \mathbf{C}_T \mathbf{K} \left[\sigma^2 (\mathbf{I} + \tilde{\delta}(\mathbf{K}) \mathbf{K}^H \mathbf{C}_T \mathbf{K}) \right]^{-1} \\ \tilde{\delta}(\mathbf{K}) &= \frac{1}{t} \text{Tr} \mathbf{C}_R \left[\sigma^2 (\mathbf{I} + \delta(\mathbf{K}) \mathbf{C}_R) \right]^{-1} \end{aligned}$$

while $(\delta(\mathbf{K}_d), \tilde{\delta}(\mathbf{K}_d))$ are the positive solutions of

$$\begin{aligned} \delta(\mathbf{K}_d) &= \frac{1}{t} \text{Tr} \mathbf{\Lambda} \left[\sigma^2 (\mathbf{I} + \tilde{\delta}(\mathbf{K}_d) \mathbf{\Lambda}) \right]^{-1} \\ \tilde{\delta}(\mathbf{K}_d) &= \frac{1}{t} \text{Tr} \mathbf{C}_R \left[\sigma^2 (\mathbf{I} + \delta(\mathbf{K}_d) \mathbf{C}_R) \right]^{-1} \end{aligned}$$

As $\mathbf{K}^H \mathbf{C}_T \mathbf{K} = \mathbf{W} \mathbf{\Lambda} \mathbf{W}^H$, for each $\tilde{\kappa} > 0$, we have

$$\frac{1}{t} \text{Tr} \mathbf{K}^H \mathbf{C}_T \mathbf{K} [\sigma^2 (\mathbf{I} + \tilde{\kappa} \mathbf{K}^H \mathbf{C}_T \mathbf{K})]^{-1} = \frac{1}{t} \text{Tr} \mathbf{\Lambda} [\sigma^2 (\mathbf{I} + \tilde{\kappa} \mathbf{\Lambda})]^{-1}$$

Therefore, $(\delta(\mathbf{K}), \tilde{\delta}(\mathbf{K}))$ and $(\delta(\mathbf{K}_d), \tilde{\delta}(\mathbf{K}_d))$ are positive solutions of the same system. The uniqueness of the solutions yields $(\delta(\mathbf{K}), \tilde{\delta}(\mathbf{K})) = (\delta(\mathbf{K}_d), \tilde{\delta}(\mathbf{K}_d))$. From this, it is easy to check that $(\gamma(\mathbf{K}), \tilde{\gamma}(\mathbf{K})) = (\gamma(\mathbf{K}_d), \tilde{\gamma}(\mathbf{K}_d))$. $\hat{I}_{mmse}(\mathbf{K})$ can thus be written as

$$\hat{I}_{mmse}(\mathbf{K}) = \sum_{j=1}^t \log \left(\frac{1}{(\mathbf{I} + \tilde{\delta}(\mathbf{K}_d) \mathbf{W} \mathbf{\Lambda} \mathbf{W}^H)_{j,j}^{-1}} \right)$$

$(\mathbf{I} + \tilde{\delta}(\mathbf{K}_d) \mathbf{W} \mathbf{\Lambda} \mathbf{W}^H)_{j,j}^{-1}$ is given by

$$(\mathbf{I} + \tilde{\delta}(\mathbf{K}_d) \mathbf{W} \mathbf{\Lambda} \mathbf{W}^H)_{j,j}^{-1} = \sum_{k=1}^t \frac{|W_{j,k}|^2}{1 + \tilde{\delta}(\mathbf{K}_d) \lambda_k}$$

where $W_{j,k}$ is the entry (j, k) of unitary matrix \mathbf{W} and where $\mathbf{\Lambda} = \text{Diag}(\lambda_1, \dots, \lambda_t)$. The function $y \rightarrow \log \frac{1}{y}$ is convex on \mathbb{R}^+ . As $\sum_{k=1}^t |W_{j,k}|^2 = 1$ (because \mathbf{W} is unitary), we have

$$\log \left[\frac{1}{\sum_{k=1}^t \frac{|W_{j,k}|^2}{1 + \tilde{\delta}(\mathbf{K}_d) \lambda_k}} \right] \leq \sum_{k=1}^t |W_{j,k}|^2 \log(1 + \tilde{\delta}(\mathbf{K}_d) \lambda_k)$$

Summing over j , and using that $\sum_{j=1}^t |W_{j,k}|^2 = 1$, we get that

$$\hat{I}_{mmse}(\mathbf{K}) \leq \sum_{k=1}^t \log(1 + \tilde{\delta}(\mathbf{K}_d) \lambda_k) = \hat{I}_{mmse}(\mathbf{K}_d)$$

We now establish that $\bar{J}_{mmse}(\mathbf{K}) \leq \bar{J}_{mmse}(\mathbf{K}_d)$.

We recall that

$$\sigma^2 \mathbf{T}_{T,j,j}(\mathbf{K}) = \sum_{k=1}^t \frac{|W_{j,k}|^2}{1 + \tilde{\delta}(\mathbf{K}_d) \lambda_k}$$

Similarly, we have

$$[(\sigma^2 \mathbf{T}_T)^2]_{j,j}(\mathbf{K}) = \sum_{k=1}^t \frac{|W_{j,k}|^2}{(1 + \tilde{\delta}(\mathbf{K}_d) \lambda_k)^2}$$

As $\sum_{k=1}^t |W_{j,k}|^2 = 1$, the convexity of function $x \rightarrow x^2$ implies that

$$[(\sigma^2 \mathbf{T}_T(\mathbf{K}))_{j,j}]^2 \leq [(\sigma^2 \mathbf{T}_T(\mathbf{K}))^2]_{j,j}$$

This implies that

$$\bar{J}_{mmse}(\mathbf{K}) \leq \frac{1}{t} \sum_{j=1}^t (1 - \sigma^2 \mathbf{T}_{T,j,j}(\mathbf{K}))^2$$

We also remark that

$$1 - \sigma^2 \mathbf{T}_{T,j,j}(\mathbf{K}) = \mathbf{w}_j(\mathbf{I} - \sigma^2 \mathbf{T}_T(\mathbf{K}_d))\mathbf{w}_j^H$$

where \mathbf{w}_j represents the row j of \mathbf{W} . As matrix $\mathbf{I} - \sigma^2 \mathbf{T}(\mathbf{K}_d)$ is diagonal and matrix \mathbf{W} is unitary, we have

$$\sum_{j=1}^t (\mathbf{w}_j(\mathbf{I} - \sigma^2 \mathbf{T}_T(\mathbf{K}_d))\mathbf{w}_j^H)^2 \leq \sum_{j=1}^t (1 - \sigma^2 \mathbf{T}_{T,j,j}(\mathbf{K}_d))^2$$

We finally note that

$$\frac{1}{2\tilde{\delta}(\mathbf{K}_d)^2} \frac{\tilde{\gamma}(\mathbf{K}_d)}{1 - \sigma^4 \gamma(\mathbf{K}_d)\tilde{\gamma}(\mathbf{K}_d)} \sum_{j=1}^t (1 - \sigma^2 \mathbf{T}_{T,j,j}(\mathbf{K}_d))^2 = \bar{J}_{mmse}(\mathbf{K}_d)$$

$\bar{J}_{mmse}(\mathbf{K}) \leq \bar{J}_{mmse}(\mathbf{K}_d)$ follows from the equalities $(\delta(\mathbf{K}), \tilde{\delta}(\mathbf{K})) = (\delta(\mathbf{K}_d), \tilde{\delta}(\mathbf{K}_d))$ and $(\gamma(\mathbf{K}), \tilde{\gamma}(\mathbf{K})) = (\gamma(\mathbf{K}_d), \tilde{\gamma}(\mathbf{K}_d))$. This completes the proof of Proposition 4.

Proposition 4 shows that there is no restriction to look an optimal precoder in the following set \mathcal{K}_d

$$\mathcal{K}_d = \{\mathbf{K} \in \mathcal{K}, \mathbf{K}^H \mathbf{C}_T \mathbf{K} \text{ diagonal}\} \quad (8.43)$$

This allows to formulate the evaluation of an optimal precoder as a t -dimensional optimization problem rather than a t^2 dimensional one. In order to state the corresponding result, we first slightly change our notations. If $\mathbf{K} \in \mathcal{K}_d$, the quantities $\delta(\mathbf{K}), \tilde{\delta}(\mathbf{K}), \dots$ are actually functions of the entries of the diagonal matrix $\mathbf{\Lambda} = \mathbf{K}^H \mathbf{C}_T \mathbf{K}$. Therefore, for $\mathbf{K} \in \mathcal{K}_d$, $\delta(\mathbf{K}), \tilde{\delta}(\mathbf{K}), \dots$ will be denoted $\delta(\mathbf{\Lambda}), \tilde{\delta}(\mathbf{\Lambda}), \dots$

The main result of this section is the following theorem.

Theorem 4. *Let $\mathbf{C}_T = \mathbf{U}\mathbf{D}\mathbf{U}^H$ be the eigenvalues/eigenvectors decomposition of matrix \mathbf{C}_T . Let $\mathbf{\Lambda}_{opt} = \text{Diag}(\lambda_{1,opt}, \dots, \lambda_{t,opt})$ be a positive diagonal matrix solution of the optimization problem*

Problem 5. *Maximize $\sum_{j=1}^t \log_2 \left(1 + \lambda_j \tilde{\delta}(\mathbf{\Lambda}) \right) + \frac{1}{2} \frac{\sigma^4 \gamma(\mathbf{\Lambda}) \tilde{\gamma}(\mathbf{\Lambda})}{1 - \sigma^4 \gamma(\mathbf{\Lambda}) \tilde{\gamma}(\mathbf{\Lambda})}$ under the constraints*

$$\mathbf{\Lambda} = \text{Diag}(\lambda_1, \dots, \lambda_t) \geq 0, \quad \frac{1}{t} \text{Tr}(\mathbf{D}^{-1} \mathbf{\Lambda}) \leq 1 \quad (8.44)$$

Then, matrix \mathbf{K}_{opt} defined by

$$\mathbf{K}_{opt} = \mathbf{U}\mathbf{D}^{-1/2}\mathbf{\Lambda}_{opt}^{1/2} \quad (8.45)$$

belongs to \mathcal{K}_d , and maximizes \bar{I}_{mmse} .

Proof. In order to prove Theorem 4, we consider a precoding matrix $\mathbf{K} \in \mathcal{K}_d$, and denote $\mathbf{\Lambda} = \text{Diag}(\lambda_1, \lambda_2, \dots, \lambda_t)$ the diagonal matrix $\mathbf{K}^H \mathbf{C}_T \mathbf{K}$. We assume that the diagonal entries $(\lambda_j)_{j=1, \dots, t}$ of $\mathbf{\Lambda}$ are arranged in decreasing order. It is clear that \mathbf{K} can be written as $\mathbf{K} = \mathbf{U}\mathbf{D}^{-1/2}\mathbf{\Theta}\mathbf{\Lambda}^{1/2}$ where $\mathbf{\Theta}$ is a unitary matrix. As $\frac{1}{t}\text{Tr}(\mathbf{K}\mathbf{K}^H)$ is supposed less than or equal to 1, matrices $\mathbf{\Lambda}$ and $\mathbf{\Theta}$ satisfy $\frac{1}{t}\text{Tr}\mathbf{D}^{-1}\mathbf{\Theta}\mathbf{\Lambda}\mathbf{\Theta}^H \leq 1$. Each precoder $\mathbf{K} \in \mathcal{K}_d$ can thus be parameterized by the unitary matrix $\mathbf{\Theta}$ and the positive diagonal matrix $\mathbf{\Lambda}$. As $\mathbf{K} \in \mathcal{K}_d$, one can check easily that $\bar{J}_{mmse}(\mathbf{\Lambda})$ reduces to $\frac{1}{2} \frac{\gamma(\mathbf{\Lambda})\tilde{\gamma}(\mathbf{\Lambda})}{1-\sigma^4\gamma(\mathbf{\Lambda})\tilde{\gamma}(\mathbf{\Lambda})}$. Therefore, Problem 5 is equivalent to the Problem

Problem 6. Maximize over $\mathbf{\Lambda}$ and $\mathbf{\Theta}$ $\sum_{j=1}^t \log_2 \left(1 + \lambda_j \tilde{\delta}(\mathbf{\Lambda}) \right) + \frac{1}{2} \frac{\gamma(\mathbf{\Lambda})\tilde{\gamma}(\mathbf{\Lambda})}{1-\sigma^4\gamma(\mathbf{\Lambda})\tilde{\gamma}(\mathbf{\Lambda})}$ under the constraints

$$\mathbf{\Lambda} = \text{Diag}(\lambda_1, \dots, \lambda_t) \geq 0, \quad \mathbf{\Theta} \text{unitary}, \quad \frac{1}{t}\text{Tr}(\mathbf{D}^{-1}\mathbf{\Theta}\mathbf{\Lambda}\mathbf{\Theta}^H) \leq 1 \quad (8.46)$$

Let $(\mathbf{\Lambda}_*, \mathbf{\Theta}_*)$ be a solution of the above problem. The diagonal elements of $\mathbf{\Lambda}_*$ and \mathbf{D} are arranged in decreasing order. Therefore (see the Appendix of [32]), the following inequality holds

$$\frac{1}{t}\text{Tr}\mathbf{D}^{-1}\mathbf{\Theta}_*\mathbf{\Lambda}_*\mathbf{\Theta}_*^H \geq \frac{1}{t}\text{Tr}\mathbf{D}^{-1}\mathbf{\Lambda}_* \quad (8.47)$$

Inequality (8.47) implies that if $(\mathbf{\Lambda}_*, \mathbf{\Theta}_*)$ is a solution of Problem 6, then, $(\mathbf{\Lambda}_*, \mathbf{I})$ is a solution of Problem 5. This shows that the optimization of \bar{I}_{mmse} is equivalent to Problem 5. This completes the proof of Theorem 4.

Remark 1. We mention that it is not obvious that the singular vectors of the precoders that optimize the true mutual information I_{mmse} have the structure (8.45). To our best knowledge, this is still an open question. One of the merit of the present asymptotic analysis is thus to show that the use of precoders (8.45) is relevant.

8.5 Maximization of \bar{I}_{mmse} .

8.5.1 Maximization of \hat{I}_{mmse} in the case of i.i.d. channels.

Problem 5 cannot in general be solved in closed form. In this paragraph, we consider the case $r = t$, $\mathbf{C}_R = \mathbf{C}_T = \mathbf{I}_t$ for which some analytical results can be obtained. We establish in particular that the transmission scheme maximizing \hat{I}_{mmse} is not a uniform power allocation along all the antennas, but an antenna selection scheme. This tends to indicate that in the context of i.i.d. channels, an antenna selection may provide higher values of I_{mmse} than a uniform power allocation over the t available transmit antennas. Therefore, even in the simplest channels context, the maximization of I_{mmse} and of the usual Shannon mutual information I are different problems.

Theorem 4 implies that precoders \mathbf{K}_{opt} maximizing \hat{I}_{mmse} can be written as $\mathbf{K}_{opt} = \mathbf{\Lambda}_{opt}^{1/2}$ where $\mathbf{\Lambda}_{opt}$ is solution of the problem

Problem 7. Maximize $\sum_{j=1}^t \log(1 + \lambda_j \tilde{\delta}(\mathbf{\Lambda}))$ under the constraints $\mathbf{\Lambda} = \text{Diag}(\lambda_1, \dots, \lambda_t) \geq 0$ and $\frac{1}{t} \text{Tr}(\mathbf{\Lambda}) \leq 1$ where $\tilde{\delta}(\mathbf{\Lambda})$ is the unique positive solution of the equation

$$\sigma^2 \tilde{\delta} + \frac{1}{t} \sum_{j=1}^t \frac{\lambda_j \tilde{\delta}}{1 + \lambda_j \tilde{\delta}} = 1 \quad (8.48)$$

It is easy to check that $\tilde{\delta}(\mathbf{\Lambda})$ is the positive solution of (8.48) in the particular context considered here. This justifies the statement of Problem 7. The solution of this problem is given in the following Proposition.

Proposition 5. The diagonal entries of the optimal matrices $\mathbf{\Lambda}_{opt}$ are either 0, either equal to $\frac{t}{s}$ where $s \leq t$, the number of non zero entries of $\mathbf{\Lambda}_{opt}$, is the integer that maximizes

$$s \log \left[\frac{\frac{t}{s} - 1 + \sigma^2 + \sqrt{(\frac{t}{s} - 1 + \sigma^2)^2 + 4\sigma^2}}{2\sigma^2} \right] \quad (8.49)$$

Proof. We first show that any optimal matrix $\mathbf{\Lambda}_{opt}$ solution of Problem 7 verifies $\frac{1}{t} \text{Tr}(\mathbf{\Lambda}_{opt}) = 1$. For this, we consider a positive diagonal matrix $\mathbf{\Lambda}$ for which $\frac{1}{t} \text{Tr}(\mathbf{\Lambda}) < 1$, and establish that if $\mathbf{\Gamma}$ is the positive diagonal matrix with normalized trace 1 defined by $\mathbf{\Gamma} = \frac{1}{\frac{1}{t} \text{Tr}(\mathbf{\Lambda})} \mathbf{\Lambda}$, then $\hat{I}_{mmse}(\mathbf{\Gamma}) > \hat{I}_{mmse}(\mathbf{\Lambda})$. For this, we show that function

$\mu \rightarrow \mu\tilde{\delta}(\mu\mathbf{\Lambda})$ is strictly increasing on \mathbb{R}^+ . We remark that $\tilde{\delta}(\mu\mathbf{\Lambda})$ is the unique positive solution of the equation

$$\sigma^2\tilde{\delta} + \frac{1}{t} \sum_{j=1}^t \frac{\lambda_j \mu \tilde{\delta}}{1 + \lambda_j \mu \tilde{\delta}} = 1 \quad (8.50)$$

or equivalently, that $\mu\tilde{\delta}(\mu\mathbf{\Lambda})$ is the unique solution of the equation $g_\mu(\tilde{\rho}) = 1$ where g_μ is defined by

$$g_\mu(\tilde{\rho}) = \frac{\sigma^2}{\mu} \tilde{\rho} + \frac{1}{t} \sum_{j=1}^t \frac{\lambda_j \tilde{\rho}}{1 + \lambda_j \tilde{\rho}}$$

For each μ , function $\tilde{\rho} \rightarrow g_\mu(\tilde{\rho})$ is strictly increasing. Moreover, if $\mu_1 < \mu_2$, then $g_{\mu_1}(\tilde{\rho}) > g_{\mu_2}(\tilde{\rho})$. From this, we get immediately that $\mu_1\tilde{\delta}(\mu_1\mathbf{\Lambda}) < \mu_2\tilde{\delta}(\mu_2\mathbf{\Lambda})$. We have thus shown that $\mu \rightarrow \mu\tilde{\delta}(\mu\mathbf{\Lambda})$ is strictly increasing. We put $\mu = \frac{1}{\frac{1}{t}\text{Tr}(\mathbf{\Lambda})}$. As $\frac{1}{t}\text{Tr}(\mathbf{\Lambda}) < 1$, μ is strictly greater than 1. Therefore, $\mu\tilde{\delta}(\mu\mathbf{\Lambda}) > \tilde{\delta}(\mathbf{\Lambda})$ or $\mu\tilde{\delta}(\mathbf{\Gamma}) > \tilde{\delta}(\mathbf{\Lambda})$. Hence,

$$\sum_{j=1}^t \log(1 + \mu \lambda_j \tilde{\delta}(\mathbf{\Gamma})) > \sum_{j=1}^t \log(1 + \lambda_j \tilde{\delta}(\mathbf{\Lambda}))$$

As the $(\mu\lambda_j)_{j=1,\dots,t}$ coincide with the diagonal entries of matrix $\mathbf{\Gamma}$, the above inequality implies that $\hat{I}_{mmse}(\mathbf{\Gamma}) > \hat{I}_{mmse}(\mathbf{\Lambda})$.

The above discussion shows that the constraint $\frac{1}{t}\text{Tr}(\mathbf{\Lambda}) \leq 1$ can be replaced by $\frac{1}{t}\text{Tr}(\mathbf{\Lambda}) = 1$ in the statement of Problem 7. In order to characterize the solutions of the maximization problem, we replace the variables $(\lambda_j)_{j=1,\dots,t}$ by the variables $(x_j)_{j=1,\dots,t}$ defined by

$$x_j = \lambda_j \tilde{\delta}(\mathbf{\Lambda}) \quad (8.51)$$

for $j = 1, \dots, t$. We claim that the maximization of \hat{I}_{mmse} over the constraints $\lambda_j \geq 0$ for $j = 1, \dots, t$ and $\frac{1}{t}\text{Tr}(\mathbf{\Lambda}) = 1$ is equivalent to the following problem

Problem 8. Maximize $\sum_{j=1}^t \log(1 + x_j)$ under the constraints $x_j \geq 0$ for each $j = 1, \dots, t$, and

$$\sigma^2 \frac{1}{t} \sum_{j=1}^t x_j + \frac{1}{t} \sum_{j=1}^t \frac{x_j}{1 + x_j} = 1 \quad (8.52)$$

Indeed, let $(x_j)_{j=1,\dots,t}$ be positive numbers satisfying (8.52), and consider $\tilde{\delta} = \frac{1}{t} \sum_{j=1}^t x_j$ and $\lambda_j = \frac{x_j}{\tilde{\delta}}$ for $j = 1, \dots, t$. The matrix $\mathbf{\Lambda} = \text{Diag}(\lambda_1, \dots, \lambda_t)$ is positive and satisfies $\frac{1}{t}\text{Tr}(\mathbf{\Lambda}) = 1$. Moreover, $\tilde{\delta}$ is solution of the equation (8.48), which implies that $\tilde{\delta} = \tilde{\delta}(\mathbf{\Lambda})$. Conversely, if $\mathbf{\Lambda} = \text{Diag}(\lambda_1, \dots, \lambda_t)$ is positive and satisfies

$\frac{1}{t}\text{Tr}(\mathbf{\Lambda}) = 1$, the $(x_j)_{j=1,\dots,t}$ defined by (8.51) are positive and satisfy the constraint (8.52). The conclusion follows from the observation that $\sum_{j=1}^t \log(1 + \lambda_j \tilde{\delta}(\mathbf{\Lambda})) = \sum_{j=1}^t \log(1 + x_j)$.

The Karush- Kuhn-Tucker (KKT) conditions provide necessary conditions for optimality. If $\mathbf{x} = (x_1, \dots, x_t)^T$ is a solution of Problem 8, then, it exists μ for which

$$\begin{aligned} \frac{\partial \mathcal{L}(\mathbf{x}, \mu)}{\partial x_j} &= 0 \text{ if } x_j > 0 \\ \frac{\partial \mathcal{L}(\mathbf{x}, \mu)}{\partial x_j} &\leq 0 \text{ if } x_j = 0 \end{aligned} \quad (8.53)$$

where $\mathcal{L}(\mathbf{x}, \mu)$ is defined by

$$\mathcal{L}(\mathbf{x}, \mu) = \sum_{j=1}^t \log(1 + x_j) - \frac{1}{\mu} \left[\sigma^2 \sum_{j=1}^t x_j + \sum_{j=1}^t \frac{x_j}{1 + x_j} \right]$$

If $x_j > 0$, we obtain

$$\sigma^2(1 + x_j) + \frac{1}{1 + x_j} = \mu$$

and if $x_j = 0$, we have

$$\mu \leq 1 + \sigma^2 \quad (8.54)$$

If $s \leq t$ is the number of non zero x_j 's, we have also

$$\mu = \sigma^2 + \sigma^2 \frac{1}{s} \sum_{j=1}^t x_j + \frac{1}{s} \sum_{x_j > 0} \frac{1}{1 + x_j} \quad (8.55)$$

or

$$\mu = \sigma^2 + \sigma^2 \frac{1}{s} \sum_{j=1}^t x_j + \frac{1}{s} \left(\sum_{j=1}^t \frac{1}{1 + x_j} - (t - s) \right)$$

Using the identity $\frac{x_j}{1+x_j} = 1 - \frac{1}{1+x_j}$, we get that the constraint (8.52) can also be written as

$$\frac{1}{t} \sum_{j=1}^t \frac{1}{1 + x_j} = \sigma^2 \frac{1}{t} \sum_{j=1}^t x_j$$

μ is therefore given by

$$\mu = \sigma^2 + 2\sigma^2 \frac{t}{s} \left(\frac{1}{t} \sum_{j=1}^t x_j \right) + 1 - \frac{t}{s} \quad (8.56)$$

We also note that $x_j > 0$ is a solution of the equation

$$\sigma^2 x_j^2 + (2\sigma^2 - \mu)x_j + (1 + \sigma^2 - \mu) = 0 \quad (8.57)$$

If $\mu > 1 + \sigma^2$, (8.54) implies that $s = t$. Moreover, the equation (8.57) has a single strictly positive solution y . Therefore, $x_j = y$ for $j = 1, \dots, t$. Using the correspondence (8.51) between \mathbf{x} and $\mathbf{\Lambda}$, we obtain immediately that $\mathbf{\Lambda} = \mathbf{I}_t$.

We now consider the case $\mu \leq 1 + \sigma^2$. If $\mu < 2\sigma$, this equation has no real solution. Therefore, μ must be greater than 2σ . The equation must have at least a positive solution. As $1 + \sigma^2 - \mu \geq 0$, this implies that $\mu > 2\sigma^2$. In sum, μ must be greater than $\max(2\sigma, 2\sigma^2)$, and the equation (8.57) has 2 positive solutions y_1 and y_2 given by

$$\begin{aligned} y_1 &= \frac{\mu - 2\sigma^2 + \sqrt{\mu^2 - 4\sigma^2}}{2\sigma^2} \\ y_2 &= \frac{\mu - 2\sigma^2 - \sqrt{\mu^2 - 4\sigma^2}}{2\sigma^2} \end{aligned}$$

Therefore, each non zero x_j can be equal to y_1 or to y_2 . We denote $\#\{j, x_j = y_1\}$ as $\frac{s}{2} + u$ and $\#\{j, x_j = y_2\}$ as $\frac{s}{2} - u$ where u is an integer if s is even and u is the sum of $1/2$ with an integer if s is odd. Note that if $(x_j)_{j=1, \dots, t}$ is a solution of Problem 8, u must be positive because $y_1 > y_2$ and

$$\sum_{j=1}^t \log(1 + x_j) = (u + \frac{s}{2}) \log(1 + y_1) + (u - \frac{s}{2}) \log(1 + y_2)$$

$\frac{1}{t} \sum_{j=1}^t x_j$ is given by

$$\frac{1}{t} \sum_{j=1}^t x_j = \frac{s}{t} \frac{\mu - 2\sigma^2}{2\sigma^2} + \frac{u}{t} \frac{\sqrt{\mu^2 - 4\sigma^2}}{\sigma^2}$$

Plugging this into (8.56) and solving the equation w.r.t. μ yields to

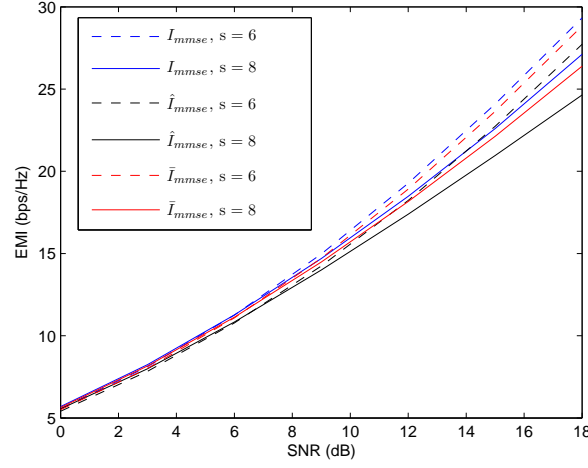
$$\mu = \sqrt{4\sigma^2 + \left(\frac{s}{2u}\sigma^2 + \frac{t-s}{2u} \right)^2} \quad (8.58)$$

This allows to express y_1 and y_2 in terms of t, s, u, σ^2 . After some calculations, we obtain that $\sum_{j=1}^t \log(1 + x_j) = (u + \frac{s}{2}) \log(1 + y_1) + (u - \frac{s}{2}) \log(1 + y_2)$ is given by

$$\sum_{j=1}^t \log(1 + x_j) = \frac{s}{2} \log \frac{1}{\sigma^2} + u \log \left[\frac{\sqrt{1 + b^2 u^2} + 1}{\sqrt{1 + b^2 u^2} - 1} \right] \quad (8.59)$$

where b^2 is defined by

$$b^2 = \frac{16\sigma^2}{(t - s + \sigma^2 s)^2}$$

FIGURE 8.3 – Relevance of the antenna selection scheme, $s = 6$ versus $s = 8$

It is easily seen that the righthandside of (8.59), considered as a function of u , is increasing on \mathbb{R}^+ . Therefore, it is maximum for $u = \frac{s}{2}$. This implies that $\#\{j, x_j = y_1\} = s$ and $\#\{j, x_j = y_2\} = 0$. Moreover, the righthandside of (8.59) for $u = s/2$ coincides with (8.49). This completes the proof of Proposition 5.

We now check numerically that for certain values of σ^2 , s does not coincide with t . In figure 8.3, we have considered the case $r = t = 8$, and represented the values of \hat{I}_{mmse} for $s = 6$ and $s = 8$. It is clear that if the SNR is greater than 8 dB, then $s = 6$ provides higher values of \hat{I}_{mmse} . The values of \bar{I}_{mmse} and I_{mmse} are still higher for $s = 6$ rather than for $s = 8$. This confirms that the antenna selection scheme may be better than the uniform power allocation across all the transmit antennas. Figure 8.4 represents $I_{mmse}, \hat{I}_{mmse}, \bar{I}_{mmse}$ versus s when the SNR is equal to 15dB, and demonstrates that $s = 6$ is the optimum value of I_{mmse} .

We note that if $s \neq t$, function \hat{I}_{mmse} reaches its maximum at different points because there are more than one diagonal matrix whose entries are either 0 either $\frac{t}{s}$. Function \hat{I}_{mmse} is thus a non concave function of the precoding matrix. Using the trick introduced in [42], it is possible to establish that I_{mmse} is itself, in general, non concave.

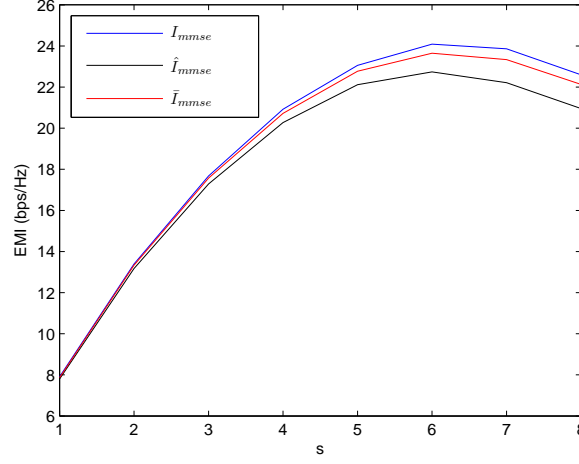


FIGURE 8.4 – Relevance of the antenna selection scheme, SNR = 15 dB

8.5.2 Study of Problem 5.

We consider again the optimization of \bar{I}_{mmse} in the bi-correlated case. Theorem 4 shows that the determination of an optimal precoder \mathbf{K}_{opt} needs to solve the optimization Problem 5. As this problem cannot be solved in closed form, we use a gradient algorithm. We propose to parameterize λ_j by $\lambda_j = \alpha_j^2$ in order to get rid of the constraint $\lambda_j \geq 0$, and to use a standard gradient algorithm with projection on the constraint $\frac{1}{t} \sum_{j=1}^t \frac{\alpha_j^2}{d_j} \leq 1$ at each iteration. Note that the convergence of this algorithm towards a global maximum of \bar{I}_{mmse} is not guaranteed because this last function is probably non concave in general.

8.5.3 Numerical illustration

We present some simulation results to illustrate the impact of the precoder optimization scheme in the case $r = t = 4$. \mathbf{C}_T and \mathbf{C}_R are generated according to model (8.38). In the present numerical experiment, $(\sigma_{\phi_T}, \phi_T) = (0.5, \frac{\pi}{4})$ and $(\sigma_{\phi_R}, \phi_R) = (0.4, \frac{\pi}{12})$.

In figure 8.5, we provide the mutual informations I_{mmse} (evaluated using Monte Carlo simulations, 1000 channel realizations are used) corresponding to the following precoding schemes :

- (i) No precoding
- (ii) Maximization of $\bar{I}_{mmse}(\mathbf{K})$ for precoders structured as in (8.45)

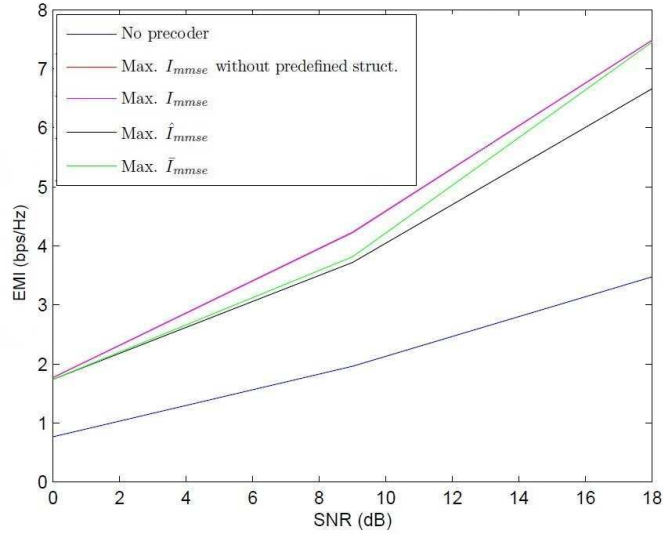


FIGURE 8.5 – Impact of precoding scheme

- (iii) Maximization of $\hat{I}_{mmse}(\mathbf{K})$ for precoders structured as in (8.45)
- (iv) Maximization of $I_{mmse}(\mathbf{K})$ for precoders structured as in (8.45)
- (v) Maximization of $I_{mmse}(\mathbf{K})$ when the precoders have no particular structure.

The various maximizations are achieved by the gradient algorithm with projection on the relevant constraint. Note that the gradients of $I_{mmse}(\mathbf{K})$ w.r.t. the parameters $(\alpha_j)_{j=1,\dots,t}$ and w.r.t. the entries of \mathbf{K} have no closed form expression. At each iteration of the algorithm, they are evaluated by Monte Carlo simulations (1000 channel realizations are used). This explains why the direct maximization of I_{mmse} leads to very high computational cost algorithms.

We now comment figure 8.5. We first compare precoding schemes (iv) and (v). The two curves match perfectly, showing that in practice, the structure (8.45) seems to optimize $I_{mmse}(\mathbf{K})$ even for $r = t = 4$. The comparison of schemes (ii) and (iii) indicates that the use of the improved approximation \bar{I}_{mmse} allows to obtain significant gains for SNRs greater than 10 dB. We finally observe that the precoding schemes (ii) and (iv,v) provide very close mutual informations when $\text{SNR} < 2$ dB and $\text{SNR} > 10$ dB. Finally, the comparison of (i) (no precoding) with the other schemes shows that the precoding allows to increase significantly I_{mmse} .

We finally compare the processing time (on a 1.83GHz Intel) needed by schemes (ii), (iii), (iv)

Precoding scheme	Processing time (s)
(ii) maximization of \bar{I}_{mmse}	0.39
(iii) maximization of \hat{I}_{mmse}	0.25
(iv) maximization of I_{mmse}	337.6

It is seen that the processing times needed to implement schemes (ii) and (iii) are almost 1000 times smaller than in the context of scheme (iv), while the use of the improved approximation \bar{I}_{mmse} instead of \hat{I}_{mmse} does not lead to a significant increase of the computational cost.

8.6 Concluding remarks.

We summarize the advantages of our asymptotic analysis of I_{mmse} . It first allows to prove the relevance of precoders $\mathbf{K} = \mathbf{U}\mathbf{D}^{-1/2}\mathbf{\Lambda}^{1/2}$, where $\mathbf{\Lambda}$ is a positive diagonal matrix. Second, the entries of the optimum matrix $\mathbf{\Lambda}$ are solution of an optimization problem that can be solved by a computationally attractive gradient algorithm. If, in contrast, matrix $\mathbf{\Lambda}$ was designed to maximize the true mutual information I_{mmse} , the corresponding gradient algorithm would have a high computational cost. This is because this function of $\mathbf{\Lambda}$, as well as its derivatives w.r.t. the entries of $\mathbf{\Lambda}$, cannot be expressed in closed form. They have to be evaluated by Monte Carlo simulations, thus complicating a lot the maximization algorithm.

Acknowledgements. The authors thank Aris Moustakas for suggesting that the relative error of the approximation \bar{I}_{mmse} was a $O(\frac{1}{t^2})$ term and not a $O(\frac{1}{t^{3/2}})$ term. Useful discussions with Walid Hachem and Jamal Najim are also acknowledged.

8.7 Proof of Proposition 3.

The proof of Proposition 3 uses extensively the Nash-Poincaré inequality as well as an integration by part formula valid in the Gaussian random matrices context. The combined use of these two tools was introduced recently by Pastur in [41] in the context of simple models. This method was used in order to evaluate the asymptotic behaviour of the Shannon capacity of bi-correlated Rayleigh MIMO channels in [16] and of bi-correlated Rician MIMO channels in [42]. We however notice that Proposition 3 has not been established in [16] and [42].

Let $\Phi(\mathbf{Y})$ be a function of the entries of matrix \mathbf{Y} defined by (8.9). Then, under certain extra assumptions on Φ (see [16]), the following Nash-Poincaré inequality holds true :

$$\text{Var}(\Phi(\mathbf{Y})) \leq \frac{1}{t} \sum_{i=1}^t \sum_{j=1}^r d_i \tilde{d}_j \mathbb{E} \left[\left| \frac{\partial \Phi(\mathbf{Y})}{\partial Y_{i,j}} \right|^2 + \left| \frac{\partial \Phi(\mathbf{Y})}{\partial \bar{Y}_{i,j}} \right|^2 \right]. \quad (8.60)$$

where $\bar{Y}_{i,j}$ represents the complex conjugate of $Y_{i,j}$. We also recall that the integration by part formula gives

$$\mathbb{E}[Y_{pq}\Phi(\mathbf{Y})] = \frac{d_p \tilde{d}_q}{t} \mathbb{E} \left[\frac{\partial \Phi(\mathbf{Y})}{\partial \bar{Y}_{pq}} \right]. \quad (8.61)$$

and

$$\mathbb{E}[\bar{Y}_{pq}\Phi(\mathbf{Y})] = \frac{d_p \tilde{d}_q}{t} \mathbb{E} \left[\frac{\partial \Phi(\mathbf{Y})}{\partial Y_{pq}} \right]. \quad (8.62)$$

We first establish (8.18). For this, we first introduce some notations. β is defined by $\beta = \frac{1}{t} \text{Tr}(\mathbf{D}\mathbf{Q})$ and $\alpha = \mathbb{E}(\beta)$. $\tilde{\mathbf{R}}$ is the $r \times r$ diagonal matrix given by

$$\tilde{\mathbf{R}} = \left[\sigma^2 (\mathbf{I}_r + \alpha \tilde{\mathbf{D}}) \right]^{-1} \quad (8.63)$$

$\tilde{\alpha}$ is defined by $\tilde{\alpha} = \frac{1}{t} \text{Tr}(\tilde{\mathbf{D}}\tilde{\mathbf{Q}})$, and \mathbf{R} is the $t \times t$ diagonal matrix given by

$$\mathbf{R} = \left[\sigma^2 (\mathbf{I}_t + \tilde{\alpha} \mathbf{D}) \right]^{-1} \quad (8.64)$$

If x is a random variable, $\overset{\circ}{x}$ represents the random variable $\overset{\circ}{x} = x - \mathbb{E}(x)$.

Using calculations similar to [16], section 4.1, we obtain that

$$\mathbb{E}((\mathbf{Q}\mathbf{y}_j)_k \bar{Y}_{i,j}) = \frac{d_i}{t} \frac{\tilde{d}_j}{1 + \alpha \tilde{d}_j} \mathbb{E}(\mathbf{Q}_{k,i}) - \frac{\tilde{d}_j}{1 + \alpha \tilde{d}_j} \mathbb{E}(\overset{\circ}{\beta}(\mathbf{Q}\mathbf{y}_j)_k \bar{Y}_{i,j})$$

for each k, i, j . Summing over j gives

$$\mathbb{E}((\mathbf{Q}\mathbf{Y}\mathbf{Y}^H)_{k,i}) = \sigma^2 d_i \tilde{\alpha} \mathbb{E}(\mathbf{Q}_{k,i}) - \sigma^2 \mathbb{E} \left(\overset{\circ}{\beta}(\mathbf{Q}\mathbf{Y}\tilde{\mathbf{D}}\tilde{\mathbf{R}}\mathbf{Y}^H)_{k,i} \right) \quad (8.65)$$

Plugging the resolvent identity (see Eq. (10) of [16])

$$\mathbf{Q}_{k,i} = \frac{\delta(k-i)}{\sigma^2} - \frac{(\mathbf{Q}\mathbf{Y}\mathbf{Y}^H)_{k,i}}{\sigma^2} \quad (8.66)$$

into (8.65), we obtain

$$\mathbb{E}(\mathbf{Q}_{k,i}) = \frac{\delta(k-i)}{\sigma^2} - d_i \tilde{\alpha} \mathbb{E}(\mathbf{Q}_{k,i}) + \mathbb{E} \left(\overset{\circ}{\beta}(\mathbf{Q}\mathbf{Y}\tilde{\mathbf{D}}\tilde{\mathbf{R}}\mathbf{Y}^H)_{k,i} \right)$$

Solving w.r.t. $\mathbb{E}(\mathbf{Q}_{k,i})$, we get

$$\mathbb{E}(\mathbf{Q}_{k,i}) = R_{i,i}\delta(k-i) + \sigma^2 \mathbb{E} \left(\overset{\circ}{\beta}(\mathbf{QY}\tilde{\mathbf{D}}\tilde{\mathbf{R}}\mathbf{Y}^H\mathbf{R})_{k,i} \right)$$

If \mathbf{u} is a deterministic unit norm row vector, we eventually obtain

$$\mathbb{E}(\mathbf{uQ}\mathbf{u}^H) = \mathbf{uR}\mathbf{u}^H + \sigma^2 \mathbb{E} \left(\overset{\circ}{\beta}\mathbf{uQY}\tilde{\mathbf{D}}\tilde{\mathbf{R}}\mathbf{Y}^H\mathbf{Ru}^H \right) \quad (8.67)$$

We now prove that the second term of the righthandside of (8.67) can be bounded by a $O(\frac{1}{t^{3/2}})$ term independent of \mathbf{u} . As $\mathbb{E}(\overset{\circ}{\beta}) = 0$, the Schwartz inequality gives

$$\left| \mathbb{E} \left(\overset{\circ}{\beta}\mathbf{uQY}\tilde{\mathbf{D}}\tilde{\mathbf{R}}\mathbf{Y}^H\mathbf{Ru}^H \right) \right| \leq \left(\mathbb{E} \left| \overset{\circ}{\beta} \right|^2 \right)^{1/2} \left[\text{Var} \left(\mathbf{uQY}\tilde{\mathbf{D}}\tilde{\mathbf{R}}\mathbf{Y}^H\mathbf{Ru}^H \right) \right]^{1/2} \quad (8.68)$$

Using the first item of (8.15) in the case $\mathbf{M} = \mathbf{D}$, we get that $\left(\mathbb{E} \left| \overset{\circ}{\beta} \right|^2 \right)^{1/2} = O(\frac{1}{t})$. In order to study the behaviour of the second term of the righthandside of (8.68), we establish the following lemma.

Lemma 2. *Let \mathbf{A} be a uniformly bounded diagonal deterministic matrix, \mathbf{u} a unit norm deterministic row vector, and \mathbf{v} a uniformly bounded deterministic row vector. Then,*

$$\text{Var}(\mathbf{uQYAY}^H\mathbf{v}^H) \leq \frac{C}{t} \quad (8.69)$$

where C is a constant independent of \mathbf{u} , \mathbf{v} , and \mathbf{A} .

Proof. In order prove the lemma, we use the Nash-Poincaré inequality (8.60) in the case $\Phi(\mathbf{Y}) = \mathbf{uQYAY}^H\mathbf{v}^H$. We define η as $\eta = \mathbf{uQYAY}^H\mathbf{v}^H$. Straightforward calculations lead to

$$\frac{\partial \eta}{\partial \bar{Y}_{i,j}} = -\mathbf{uQy}_j(\mathbf{QYAY}^H\mathbf{v}^H)_i + A_{j,j}v_i \mathbf{uQy}_j \quad (8.70)$$

We establish that

$$\sum_{i,j} \mathbb{E} \left| \frac{\partial \eta}{\partial \bar{Y}_{i,j}} \right|^2 \leq C \quad (8.71)$$

where C is a constant independent of \mathbf{u} , \mathbf{v} , \mathbf{A} . (8.70), $|A_{j,j}| \leq \|\mathbf{A}\|$ and the Schwartz inequality imply that

$$\mathbb{E} \left| \frac{\partial \eta}{\partial \bar{Y}_{i,j}} \right|^2 \leq 2|v_i|^2 \|\mathbf{A}\|^2 \mathbb{E} |\mathbf{uQy}_j|^2 + 2\mathbb{E} (|(\mathbf{QYAY}^H\mathbf{v}^H)_i|^2 |\mathbf{uQy}_j|^2)$$

Summing over i, j yields

$$\sum_{i,j} \mathbb{E} \left| \frac{\partial \eta}{\partial Y_{i,j}} \right|^2 \leq 2\|\mathbf{A}\|^2 \|\mathbf{v}\|^2 \mathbb{E} \|\mathbf{uQY}\|^2 + 2\mathbb{E} (\|\mathbf{QYAY}^H \mathbf{v}^H\|^2 \|\mathbf{uQY}\|^2)$$

$\mathbb{E}(\|\mathbf{uQY}\|^2) = \mathbb{E}(\mathbf{uQYY}^H \mathbf{Qu}^H)$. Using the resolvent identity (8.66), we obtain that $\mathbf{QYY}^H = \mathbf{I} - \sigma^2 \mathbf{Q}$. Therefore, $\mathbf{QYY}^H \mathbf{Q} = \mathbf{Q} - \sigma^2 \mathbf{Q}\mathbf{Q}$ and $\mathbf{QYY}^H \mathbf{Q} \leq \mathbf{Q}$. This implies that $\|\mathbf{uQY}\|^2 \leq \mathbf{uQ}\mathbf{u}^H$. As matrix \mathbf{Q} satisfies $\mathbf{Q} \leq \frac{\mathbf{I}}{\sigma^2}$, we obtain that

$$\|\mathbf{uQY}\|^2 \leq \frac{1}{\sigma^2} \quad (8.72)$$

In order to prove (8.71), it is thus sufficient to verify that $\mathbb{E}(\|\mathbf{QYAY}^H \mathbf{v}^H\|^2) \leq C$ where C is a constant independent of \mathbf{v} and t . For this, we remark that

$$\|\mathbf{QYAY}^H \mathbf{v}^H\|^2 \leq \frac{\mathbf{vYA}^H \mathbf{Y}^H \mathbf{YAY}^H \mathbf{v}^H}{\sigma^4}$$

A straightforward but tedious calculation gives

$$\mathbb{E}(\mathbf{vYA}^H \mathbf{Y}^H \mathbf{YAY}^H \mathbf{v}^H) = \mathbf{vD}^2 \mathbf{v}^H \left| \frac{1}{t} \text{Tr}(\mathbf{A}^H \tilde{\mathbf{D}}) \right|^2 + \mathbf{vD} \mathbf{v}^H \frac{1}{t} \text{Tr} \mathbf{D} \frac{1}{t} \text{Tr}(\mathbf{A}^H \tilde{\mathbf{D}} \mathbf{A} \tilde{\mathbf{D}})$$

As matrices \mathbf{A} , \mathbf{D} , $\tilde{\mathbf{D}}$ and vector \mathbf{v} are uniformly bounded, we obtain that $\mathbb{E}(\|\mathbf{QYAY}^H \mathbf{u}^H\|^2) \leq C$. This, in turn, implies (8.71). One can show similarly that

$$\sum_{i,j} \mathbb{E} \left| \frac{\partial \eta}{\partial Y_{i,j}} \right|^2 \leq C$$

As the $(d_i)_{i=1,\dots,t}$ and the $(\tilde{d}_j)_{j=1,\dots,r}$ are uniformly bounded (see (8.14)), (8.60) provides immediately (8.69). Lemma 2 is thus established.

(8.67) and (8.68) imply that

$$|\mathbf{u}(\mathbb{E}(\mathbf{Q}) - \mathbf{R})\mathbf{u}^H| \leq \frac{C}{t^{3/2}} \quad (8.73)$$

In order to complete the proof of (8.18), we use Theorem 3 of [16], and obtain that

$$\frac{1}{t} \text{Tr}(\tilde{\mathbf{D}}\tilde{\mathbf{R}}) = \frac{1}{t} \text{Tr}(\tilde{\mathbf{D}}\tilde{\mathbf{T}}) + O\left(\frac{1}{t^2}\right)$$

or $\tilde{\alpha} = \tilde{\delta} + O(\frac{1}{t^2})$. It is easy to check that

$$|R_{i,i} - T_{i,i}| \leq \frac{d_{max}}{\sigma^2} |\tilde{\alpha} - \tilde{\delta}|$$

Therefore, $\max_i |R_{i,i} - T_{i,i}| \leq \frac{C}{t^2}$, and $|\mathbf{u}(\mathbf{R} - \mathbf{T})\mathbf{u}^H| \leq \frac{C}{t^2}$. Using (8.73), we eventually get (8.18).

We now establish (8.20). For this, we first prove the following lemma.

Lemma 3.

$$\mathbb{E}(\mathbf{Q}_{k,i} \mathring{\mathbf{Q}}_{k',i'}) = \frac{1}{t} \frac{1}{1 + \tilde{\alpha} d_i} \mathbb{E} \left[(\mathbf{Q} \mathbf{D} \mathbf{Q})_{k,i'} (\mathbf{Q} \mathbf{Y} \tilde{\mathbf{D}} \tilde{\mathbf{R}} \mathbf{Y}^H)_{k',i} \right] + \mathbb{E} \left[\mathring{\beta} (\mathbf{Q} \mathbf{Y} \tilde{\mathbf{D}} \tilde{\mathbf{R}} \mathbf{Y}^H)_{k,i} \frac{1}{1 + \tilde{\alpha} d_i} \mathring{\mathbf{Q}}_{k',i'} \right] \quad (8.74)$$

Proof. We first note that (8.66) yields

$$\mathbb{E}(\mathbf{Q}_{k,i} \mathring{\mathbf{Q}}_{k',i'}) = -\frac{1}{\sigma^2} \mathbb{E} \left((\mathbf{Q} \mathbf{Y} \mathbf{Y}^H)_{k,i} \mathring{\mathbf{Q}}_{k',i'} \right) \quad (8.75)$$

In order to be able to express $\mathbb{E} \left((\mathbf{Q} \mathbf{Y} \mathbf{Y}^H)_{k,i} \mathring{\mathbf{Q}}_{k',i'} \right)$, we evaluate

$$\mathbb{E} \left((\mathbf{Q} \mathbf{y}_j)_k \bar{Y}_{i,j} \mathring{\mathbf{Q}}_{k',i'} \right) = \sum_{p=1}^t \mathbb{E}(\mathbf{Q}_{k,p} Y_{p,j} \bar{Y}_{i,j} \mathring{\mathbf{Q}}_{k',i'})$$

For this, we use the integration by parts formula (8.61) in the case $\Phi(\mathbf{Y}) = \mathbf{Q}_{k,p} \bar{Y}_{i,j} \mathring{\mathbf{Q}}_{k',i'}$, and obtain

$$\begin{aligned} \mathbb{E}(\mathbf{Q}_{k,p} Y_{p,j} \bar{Y}_{i,j} \mathring{\mathbf{Q}}_{k',i'}) &= \delta(p-i) \frac{d_p \tilde{d}_j}{t} \mathbb{E}(\mathbf{Q}_{k,p} \mathring{\mathbf{Q}}_{k',i'}) - \\ &\quad \frac{d_p \tilde{d}_j}{t} \mathbb{E}(\mathbf{Q}_{p,p} (\mathbf{Q} \mathbf{y}_j)_k \mathring{\mathbf{Q}}_{k',i'} \bar{Y}_{i,j}) - \frac{d_p \tilde{d}_j}{t} \mathbb{E}(\mathbf{Q}_{k,p} \mathbf{Q}_{p,i'} (\mathbf{Q} \mathbf{y}_j)_{k'} \bar{Y}_{i,j}) \end{aligned} \quad (8.76)$$

Summing over p , and expressing $\beta = \frac{1}{t} \text{Tr}(\mathbf{D} \mathbf{Q})$ as $\beta = \alpha + \mathring{\beta}$ provides

$$\begin{aligned} \mathbb{E} \left[(\mathbf{Q} \mathbf{y}_j)_k \bar{Y}_{i,j} \mathring{\mathbf{Q}}_{k',i'} \right] &= \frac{d_i \tilde{d}_j}{t} \mathbb{E}(\mathbf{Q}_{k,i} \mathring{\mathbf{Q}}_{k',i'}) - \frac{\tilde{d}_j}{t} \mathbb{E} \left[(\mathbf{Q} \mathbf{D} \mathbf{Q})_{k,i'} (\mathbf{Q} \mathbf{y}_j)_{k'} \bar{Y}_{i,j} \right] - \\ &\quad \alpha \tilde{d}_j \mathbb{E} \left[(\mathbf{Q} \mathbf{y}_j)_k \bar{Y}_{i,j} \mathring{\mathbf{Q}}_{k',i'} \right] - \tilde{d}_j \mathbb{E} \left[\mathring{\beta} (\mathbf{Q} \mathbf{y}_j)_k \bar{Y}_{i,j} \mathring{\mathbf{Q}}_{k',i'} \right] \end{aligned} \quad (8.77)$$

Solving w.r.t. $\mathbb{E} \left[(\mathbf{Q} \mathbf{y}_j)_k \bar{Y}_{i,j} \mathring{\mathbf{Q}}_{k',i'} \right]$ and summing over j gives

$$\begin{aligned} \mathbb{E} \left((\mathbf{Q} \mathbf{Y} \mathbf{Y}^H)_{k,i} \mathring{\mathbf{Q}}_{k',i'} \right) &= \sigma^2 d_i \mathbb{E}(\mathbf{Q}_{k,i} \mathring{\mathbf{Q}}_{k',i'}) - \\ &\quad \frac{\sigma^2}{t} \mathbb{E} \left[(\mathbf{Q} \mathbf{D} \mathbf{Q})_{k,i'} (\mathbf{Q} \mathbf{Y} \tilde{\mathbf{D}} \tilde{\mathbf{R}} \mathbf{Y}^H)_{k',i} \right] - \sigma^2 \mathbb{E} \left[\mathring{\beta} \mathring{\mathbf{Q}}_{k',i'} (\mathbf{Q} \mathbf{Y} \tilde{\mathbf{D}} \tilde{\mathbf{R}} \mathbf{Y}^H)_{k,i} \right] \end{aligned} \quad (8.78)$$

Plugging (8.75) into (8.78) and solving w.r.t. $\mathbb{E}(\mathbf{Q}_{k,i} \overset{\circ}{\mathbf{Q}}_{k',i'})$ gives (8.74).

We define η by $\eta = \mathbf{u}\mathbf{Q}\mathbf{u}^H$. (8.74) yields immediately

$$\mathbb{E}(\overset{\circ}{\eta})^2 = \mathbb{E}(\eta\overset{\circ}{\eta}) = \frac{\sigma^2}{t} \mathbb{E} \left(\mathbf{u}\mathbf{Q}\mathbf{D}\mathbf{Q}\mathbf{u}^H \mathbf{u}\mathbf{Q}\mathbf{Y}\tilde{\mathbf{D}}\tilde{\mathbf{R}}\mathbf{Y}^H \mathbf{R}\mathbf{u}^H \right) + \sigma^2 \mathbb{E} \left[\overset{\circ}{\beta}\overset{\circ}{\eta} \mathbf{u}\mathbf{Q}\mathbf{Y}\tilde{\mathbf{D}}\tilde{\mathbf{R}}\mathbf{Y}^H \mathbf{R}\mathbf{u}^H \right] \quad (8.79)$$

We define ρ_1 and ρ_2 by $\rho_1 = \mathbf{u}\mathbf{Q}\mathbf{D}\mathbf{Q}\mathbf{u}^H$ and $\rho_2 = \mathbf{u}\mathbf{Q}\mathbf{Y}\tilde{\mathbf{D}}\tilde{\mathbf{R}}\mathbf{Y}^H \mathbf{R}\mathbf{u}^H$. The term $\mathbb{E} \left(\mathbf{u}\mathbf{Q}\mathbf{D}\mathbf{Q}\mathbf{u}^H \mathbf{u}\mathbf{Q}\mathbf{Y}\tilde{\mathbf{D}}\tilde{\mathbf{R}}\mathbf{Y}^H \mathbf{R}\mathbf{u}^H \right)$ is given by

$$\mathbb{E} \left(\mathbf{u}\mathbf{Q}\mathbf{D}\mathbf{Q}\mathbf{u}^H \mathbf{u}\mathbf{Q}\mathbf{Y}\tilde{\mathbf{D}}\tilde{\mathbf{R}}\mathbf{Y}^H \mathbf{R}\mathbf{u}^H \right) = \mathbb{E}(\rho_1)\mathbb{E}(\rho_2) + \mathbb{E}(\overset{\circ}{\rho}_1\overset{\circ}{\rho}_2)$$

In order to evaluate $\mathbb{E}(\rho_i)$, $i = 1, 2$, we state the following Lemma

Lemma 4.

$$\left| \mathbb{E}(\mathbf{Q}\mathbf{D}\mathbf{Q})_{k,i} - \frac{d_i T_{k,i}}{1 - \sigma^4 \gamma \tilde{\gamma}} \right| < \frac{C}{t} \quad (8.80)$$

Let \mathbf{A} be a uniformly bounded diagonal deterministic matrix. Then,

$$\left| \mathbb{E}(\mathbf{Q}\mathbf{Y}\mathbf{A}\mathbf{Y}^H)_{k,i} - \sigma^2 d_i \frac{1}{t} \text{Tr}(\mathbf{A}\tilde{\mathbf{T}}\tilde{\mathbf{D}}) T_{k,i} \right| < \frac{C}{t} \quad (8.81)$$

The proof uses again the resolvent identity (8.66), the Nash-Poincaré inequality, the integration by parts formula, Theorem 3 in [16], and is omitted.

Using Lemma (4), we get that

$$\left| \sigma^2 \mathbb{E}(\rho_1)\mathbb{E}(\rho_2) - \frac{\sigma^4 \tilde{\gamma}}{1 - \sigma^4 \gamma \tilde{\gamma}} (\mathbf{u}\mathbf{T}^2 \mathbf{D}\mathbf{u}^H)^2 \right| \leq \frac{C}{t}$$

We verify that $\mathbb{E}(\overset{\circ}{\rho}_1\overset{\circ}{\rho}_2)$ is a $O(\frac{1}{t^{1/2}})$ term. We first remark that, as $\rho_1 \leq \frac{d_{max}}{\sigma^4}$, then $|\overset{\circ}{\rho}_1| \leq 2\frac{d_{max}}{\sigma^4}$. The Schwartz inequality gives

$$|\mathbb{E}(\overset{\circ}{\rho}_1\overset{\circ}{\rho}_2)| \leq 2 \frac{d_{max}}{\sigma^4} \left(\mathbb{E}|\overset{\circ}{\rho}_2|^2 \right)^{1/2} \leq \frac{C}{t^{1/2}}$$

by Lemma 2. Finally, we show that $\mathbb{E}(\overset{\circ}{\beta}\overset{\circ}{\eta}\rho_2)$ is a $O(\frac{1}{t^{3/2}})$ term. We express this term as

$$\mathbb{E}(\overset{\circ}{\beta}\overset{\circ}{\eta}\rho_2) = \mathbb{E}(\overset{\circ}{\beta}\overset{\circ}{\eta})\mathbb{E}(\rho_2) + \mathbb{E}(\overset{\circ}{\beta}\overset{\circ}{\eta}\overset{\circ}{\rho}_2)$$

Lemma 4 implies that $\mathbb{E}(\rho_2)$ is uniformly bounded, while the Schwartz inequality gives $\mathbb{E}(\overset{\circ}{\beta}\overset{\circ}{\eta}) = O(\frac{1}{t^{3/2}})$. In order to evaluate $\mathbb{E}(\overset{\circ}{\beta}\overset{\circ}{\eta}\overset{\circ}{\rho}_2)$, we write

$$\mathbb{E}(\overset{\circ}{\beta}\overset{\circ}{\eta}\overset{\circ}{\rho}_2) = \mathbb{E}(\overset{\circ}{\beta}\overset{\circ}{\eta}\overset{\circ}{\rho}_2) - \mathbb{E}(\eta)\mathbb{E}(\overset{\circ}{\beta}\overset{\circ}{\rho}_2)$$

As $\eta \leq \frac{d_{max}}{\sigma^2}$, the Schwartz inequality gives immediately that

$$\mathbb{E}(\beta \overset{\circ}{\eta} \overset{\circ}{\rho}_2) = O(\frac{1}{t^{3/2}}).$$

Putting all the pieces together completes the proof of (8.20).

In order to establish (8.21), we first need to prove the following lemma. This lemma will also be useful to establish Lemma 1 below.

Lemma 5. *Let \mathbf{M} be a uniformly bounded deterministic matrix. Then,*

$$\mathbb{E} \left| \frac{1}{t} \text{Tr} \mathbf{M} (\mathbf{Q} - \mathbb{E}(\mathbf{Q})) \right|^4 = O(\frac{1}{t^4}) \quad (8.82)$$

Moreover,

$$\sup_{u, \|u\|=1} \mathbb{E} |\mathbf{u} (\mathbf{Q} - \mathbb{E}(\mathbf{Q})) \mathbf{u}^H|^8 \leq \frac{C}{t^4} \quad (8.83)$$

We denote by ρ the random variable $\rho = \frac{1}{t} \text{Tr} \mathbf{M} (\mathbf{Q} - \mathbb{E}(\mathbf{Q}))$. $\mathbb{E}|\rho|^4$ can be written as

$$\mathbb{E}|\rho|^4 = (\mathbb{E}|\rho|^2)^2 + \text{Var}(\rho^2)$$

The first item of (8.15) implies that $(\mathbb{E}|\rho|^2)^2 = O(\frac{1}{t^4})$. In order to evaluate $\text{Var}(\rho^2)$, we use the Nash-Poincaré inequality in the case $\Phi(\mathbf{Y}) = \rho^2$.

$$\begin{aligned} \frac{\partial \rho^2}{\partial \bar{Y}_{i,j}} &= 2\rho \frac{1}{t} \sum_{p,q} \frac{\partial \mathbf{Q}_{p,q}}{\partial \bar{Y}_{i,j}} \mathbf{M}_{q,p} \\ &= -2\rho \frac{1}{t} \sum_{p,q} \mathbf{Q}_{i,q} (\mathbf{Q} \mathbf{y}_j)_p \mathbf{M}_{q,p} \\ &= -2\rho \frac{1}{t} \sum_p (\mathbf{Q} \mathbf{M})_{i,p} (\mathbf{Q} \mathbf{y}_j)_p \\ &= -2\rho \frac{1}{t} (\mathbf{Q} \mathbf{M} \mathbf{Q} \mathbf{y}_j)_i \end{aligned}$$

Therefore,

$$\sum_{i,j} \mathbb{E} \left| \frac{\partial \rho^2}{\partial \bar{Y}_{i,j}} \right|^2 = \frac{4}{t^2} \mathbb{E} (|\rho|^2 \text{Tr}(\mathbf{Y}^H \mathbf{Q} \mathbf{M}^H \mathbf{Q}^2 \mathbf{M} \mathbf{Q} \mathbf{Y}))$$

Matrix $\mathbf{M}^H \mathbf{Q}^2 \mathbf{M}$ is uniformly bounded. Therefore,

$$\text{Tr}(\mathbf{Y}^H \mathbf{Q} \mathbf{M}^H \mathbf{Q}^2 \mathbf{M} \mathbf{Q} \mathbf{Y}) \leq C \text{Tr} \mathbf{Y}^H \mathbf{Q} \mathbf{Q} \mathbf{Y} = C \text{Tr}(\mathbf{Q} \mathbf{Q} \mathbf{Y} \mathbf{Y}^H)$$

As $\mathbf{Q} \mathbf{Y} \mathbf{Y}^H = \mathbf{I} - \sigma^2 \mathbf{Q} \leq \mathbf{I}$, we obtain that

$$\text{Tr}(\mathbf{Y}^H \mathbf{Q} \mathbf{M}^H \mathbf{Q}^2 \mathbf{M} \mathbf{Q} \mathbf{Y}) \leq C \text{Tr}(\mathbf{Q})$$

Hence,

$$\sum_{i,j} \mathbb{E} \left| \frac{\partial \rho^2}{\partial \bar{Y}_{i,j}} \right|^2 \leq \frac{C}{t} \mathbb{E} \left(|\rho|^2 \frac{1}{t} \text{Tr}(\mathbf{Q}) \right) \leq \frac{C}{t} \mathbb{E}(|\rho|^2)$$

As $\mathbb{E}(|\rho|^2) = O(\frac{1}{t^2})$, this implies that

$$\sum_{i,j} \mathbb{E} \left| \frac{\partial \rho^2}{\partial \bar{Y}_{i,j}} \right|^2 = O(\frac{1}{t^3})$$

We obtain similarly that

$$\sum_{i,j} \mathbb{E} \left| \frac{\partial \rho^2}{\partial Y_{i,j}} \right|^2 = O(\frac{1}{t^3})$$

(8.82) follows immediately from the Nash-Poincaré identity.

In order to prove (8.83), we first establish that

$$\sup_{u, \|u\|=1} \mathbb{E} |\mathbf{u}(\mathbf{Q} - \mathbb{E}(\mathbf{Q})) \mathbf{u}^H|^4 \leq \frac{C}{t^2} \quad (8.84)$$

and

$$\sup_{u, \|u\|=1} \mathbb{E} |\mathbf{u}(\mathbf{Q} - \mathbb{E}(\mathbf{Q})) \mathbf{u}^H|^6 \leq \frac{C}{t^3} \quad (8.85)$$

We consider a deterministic unit norm row vector \mathbf{u} and denote by η the term $\eta = \mathbf{u}(\mathbf{Q} - \mathbb{E}(\mathbf{Q})) \mathbf{u}^H$. $\mathbb{E}|\eta|^4 = (\mathbb{E}|\eta|^2)^2 + \text{Var}(\eta^2)$. (8.20) implies that $(\mathbb{E}|\eta|^2)^2 \leq \frac{C}{t^2}$ where C is a constant which does not depend on t and \mathbf{u} . In order to evaluate the term $\text{Var}(\eta^2)$, we use the Nash-Poincaré inequality in the case $\Phi(\mathbf{Y}) = \eta^2$.

$$\frac{\partial \eta^2}{\partial \bar{Y}_{i,j}} = -2\eta \mathbf{u} \mathbf{Q} \mathbf{y}_j (\mathbf{Q} \mathbf{u}^H)_i$$

Therefore,

$$\sum_{i,j} \mathbb{E} \left| \frac{\partial \eta^2}{\partial \bar{Y}_{i,j}} \right|^2 = 4 \mathbb{E} (|\eta|^2 \mathbf{u} \mathbf{Q} \mathbf{Y} \mathbf{Y}^H \mathbf{Q} \mathbf{u}^H \mathbf{u} \mathbf{Q}^2 \mathbf{u}^H)$$

(8.72), $\mathbf{Q} \leq \frac{\mathbf{I}}{\sigma^2}$, and $\mathbb{E}|\eta|^2 \leq \frac{C}{t}$ imply that

$$\sum_{i,j} \mathbb{E} \left| \frac{\partial \eta^2}{\partial \bar{Y}_{i,j}} \right|^2 \leq \frac{C}{t}$$

We obtain similarly that

$$\sum_{i,j} \mathbb{E} \left| \frac{\partial \eta^2}{\partial Y_{i,j}} \right|^2 \leq \frac{C}{t}$$

The Nash-Poincaré inequality eventually gives $\text{Var}(\eta^2) \leq \frac{C}{t^2}$. We have therefore proved (8.84). In order to establish (8.85), we write $\mathbb{E}|\eta|^6 = (\mathbb{E}|\eta|^3)^2 + \text{Var}(\eta^3)$. The Holder inequality and (8.84) imply that $(\mathbb{E}|\eta|^3)^2 \leq \frac{C}{t^3}$. The term $\text{Var}(\eta^3)$ is also evaluated using the Nash-Poincaré inequality.

$$\frac{\partial \eta^3}{\partial \bar{Y}_{i,j}} = -3\eta^2 \mathbf{u} \mathbf{Q} \mathbf{Y}_j (\mathbf{Q} \mathbf{u}^H)_i$$

and

$$\sum_{i,j} \mathbb{E} \left| \frac{\partial \eta^3}{\partial \bar{Y}_{i,j}} \right|^2 = 9 \mathbb{E} (|\eta|^4 \mathbf{u} \mathbf{Q} \mathbf{Y} \mathbf{Y}^H \mathbf{Q} \mathbf{u}^H \mathbf{u} \mathbf{Q}^2 \mathbf{u}^H)$$

As $\mathbf{u} \mathbf{Q} \mathbf{Y} \mathbf{Y}^H \mathbf{Q} \mathbf{u}^H$ and $\mathbf{u} \mathbf{Q}^2 \mathbf{u}^H$ are uniformly bounded, (8.84) implies that

$$\sum_{i,j} \mathbb{E} \left| \frac{\partial \eta^3}{\partial \bar{Y}_{i,j}} \right|^2 \leq \frac{C}{t^2}$$

Similarly,

$$\sum_{i,j} \mathbb{E} \left| \frac{\partial \eta^3}{\partial Y_{i,j}} \right|^2 \leq \frac{C}{t^2}$$

(8.85) follows immediately from the Nash-Poincaré inequality.

Starting from $\mathbb{E}|\eta|^8 = (\mathbb{E}|\eta|^4)^2 + \text{Var}(\eta^4)$, (8.83) is proved similarly.

In order to establish (8.21), we introduce the following notations :

$$\rho_{1,k} = \mathbf{v}_k \mathbf{Q} \mathbf{D} \mathbf{Q} \mathbf{v}_k^H, \rho_{2,k} = \mathbf{v}_k \mathbf{Q} \mathbf{Y} \tilde{\mathbf{D}} \tilde{\mathbf{R}} \mathbf{Y}^H \mathbf{R} \mathbf{v}_k^H, \eta_k = \mathbf{v}_k \mathbf{Q} \mathbf{v}_k^H$$

Using (8.79) and Lemma 4, it is easy to check that

$$\sum_k \text{Var}(\kappa_k \eta_k) - \frac{1}{t} \frac{\sigma^4 \tilde{\gamma}}{1 - \sigma^4 \gamma \tilde{\gamma}} \sum_k (\kappa_k \mathbf{v}_k \mathbf{T}^2 \mathbf{D} \mathbf{v}_k^H)^2 = \mathbb{E} \left(\overset{\circ}{\beta} \left(\sum_k \kappa_k \overset{\circ}{\eta}_k \rho_{2,k} \right) \right) + O\left(\frac{1}{t}\right)$$

It therefore remains to show that

$$\mathbb{E} \left(\overset{\circ}{\beta} \left(\sum_k \kappa_k \overset{\circ}{\eta}_k \rho_{2,k} \right) \right) = O\left(\frac{1}{t}\right) \quad (8.86)$$

For this, we write $\rho_{2,k} = \overset{\circ}{\rho}_{2,k} + \mathbb{E}(\rho_{2,k})$. Therefore,

$$\mathbb{E} \left(\overset{\circ}{\beta} \left(\sum_k \kappa_k \overset{\circ}{\eta}_k \rho_{2,k} \right) \right) = \mathbb{E} \left[\overset{\circ}{\beta} \left(\sum_k \kappa_k \mathbb{E}(\rho_{2,k}) \overset{\circ}{\eta}_k \right) \right] + \mathbb{E} \left[\overset{\circ}{\beta} \left(\sum_k \kappa_k \overset{\circ}{\rho}_{2,k} \overset{\circ}{\eta}_k \right) \right] \quad (8.87)$$

The term $\mathbb{E} \left[\overset{\circ}{\beta} \left(\sum_k \kappa_k \mathbb{E}(\rho_{2,k}) \overset{\circ}{\eta}_k \right) \right]$ can also be written as $\mathbb{E} \left(\overset{\circ}{\beta} \text{Tr}(\mathbf{M}\overset{\circ}{\mathbf{Q}}) \right)$ where \mathbf{M} is the deterministic matrix defined by

$$\mathbf{M} = \sum_k \kappa_k \mathbb{E}(\rho_{2,k}) \mathbf{v}_k^H \mathbf{v}_k$$

Lemma 4 implies that $\sup_k |\mathbb{E}(\rho_{2,k})| < C$. Therefore, matrix \mathbf{M} is uniformly bounded. The first item of (8.15) thus implies that $\mathbb{E} \left| \text{Tr}(\mathbf{M}\overset{\circ}{\mathbf{Q}}) \right|^2 = O(1)$. Similarly, $\mathbb{E} |\overset{\circ}{\beta}|^2 = O(\frac{1}{t^2})$ holds. The Schwartz inequality shows that $\mathbb{E} \left(\overset{\circ}{\beta} \text{Tr}(\mathbf{M}\overset{\circ}{\mathbf{Q}}) \right) = O(\frac{1}{t})$.

In order to evaluate the second term of the righthandside of (8.87), we remark that

$$|\mathbb{E}(\overset{\circ}{\beta} \overset{\circ}{\eta}_k \overset{\circ}{\rho}_{2,k})| \leq (\mathbb{E} |\overset{\circ}{\rho}_{2,k}|^2)^{1/2} (\mathbb{E} |\overset{\circ}{\beta}_k|^4)^{1/4} (\mathbb{E} |\overset{\circ}{\eta}_k|^4)^{1/4}$$

Lemma 2 implies that $(\mathbb{E} |\overset{\circ}{\rho}_{2,k}|^2)^{1/2} = O(\frac{1}{t^{1/2}})$, and (8.84) gives $(\mathbb{E} |\overset{\circ}{\eta}_k|^4)^{1/4} = O(\frac{1}{t^{1/2}})$. As $(\mathbb{E} |\overset{\circ}{\beta}_k|^4)^{1/4} = O(\frac{1}{t})$ by (8.82), we get that

$$\sup_k |\mathbb{E}(\overset{\circ}{\beta} \overset{\circ}{\eta}_k \overset{\circ}{\rho}_{2,k})| = O(\frac{1}{t^2})$$

This, in turn, implies that the second term of the righthandside of (8.87) is a $O(\frac{1}{t})$ term. This completes the proof of (8.21).

We finally prove (8.19). We just sketch the proof because similar arguments have been used in order to establish Lemma 3. We evaluate $\mathbb{E}(\overset{\circ}{\mathbf{Q}}_{k_1, i_1} \overset{\circ}{\mathbf{Q}}_{k_2, i_2} \overset{\circ}{\mathbf{Q}}_{k_3, i_3})$ for each integers $(i_1, k_1, i_2, k_2, i_3, k_3)$. We first calculate $\mathbb{E}(\mathbf{Q}_{k_1, i_1} \overset{\circ}{\mathbf{Q}}_{k_2, i_2} \overset{\circ}{\mathbf{Q}}_{k_3, i_3})$. For this, we use the resolvent identity (8.66) and write

$$\mathbb{E}(\mathbf{Q}_{k_1, i_1} \overset{\circ}{\mathbf{Q}}_{k_2, i_2} \overset{\circ}{\mathbf{Q}}_{k_3, i_3}) = \frac{\delta(k_1 - i_1)}{\sigma^2} \mathbb{E}(\overset{\circ}{\mathbf{Q}}_{k_2, i_2} \overset{\circ}{\mathbf{Q}}_{k_3, i_3}) - \frac{1}{\sigma^2} \mathbb{E} \left[(\mathbf{Q} \mathbf{Y} \mathbf{Y}^H)_{k_1, i_1} \overset{\circ}{\mathbf{Q}}_{k_2, i_2} \overset{\circ}{\mathbf{Q}}_{k_3, i_3} \right]$$

Using the integration by parts formula as in the proof of Lemma 3, we obtain that

$$\begin{aligned} \frac{1}{\sigma^2} \mathbb{E} \left[(\mathbf{Q} \mathbf{Y} \mathbf{Y}^H)_{k_1, i_1} \overset{\circ}{\mathbf{Q}}_{k_2, i_2} \overset{\circ}{\mathbf{Q}}_{k_3, i_3} \right] &= \tilde{\alpha} d_{i_1} \mathbb{E}(\mathbf{Q}_{k_1, i_1} \overset{\circ}{\mathbf{Q}}_{k_2, i_2} \overset{\circ}{\mathbf{Q}}_{k_3, i_3}) \\ &\quad - \frac{1}{t} \mathbb{E} \left[(\mathbf{Q} \mathbf{D} \mathbf{Q})_{k_1, i_2} (\mathbf{Q} \mathbf{Y} \tilde{\mathbf{D}} \tilde{\mathbf{R}} \mathbf{Y}^H)_{k_2, i_1} \overset{\circ}{\mathbf{Q}}_{k_3, i_3} \right] \\ &\quad - \frac{1}{t} \mathbb{E} \left[(\mathbf{Q} \mathbf{D} \mathbf{Q})_{k_1, i_3} (\mathbf{Q} \mathbf{Y} \tilde{\mathbf{D}} \tilde{\mathbf{R}} \mathbf{Y}^H)_{k_3, i_1} \overset{\circ}{\mathbf{Q}}_{k_2, i_2} \right] \\ &\quad - \mathbb{E} \left[\overset{\circ}{\beta} (\mathbf{Q} \mathbf{Y} \tilde{\mathbf{D}} \tilde{\mathbf{R}} \mathbf{Y}^H)_{k_i, i_1} \overset{\circ}{\mathbf{Q}}_{k_2, i_2} \overset{\circ}{\mathbf{Q}}_{k_3, i_3} \right] \end{aligned}$$

Plugging (8.66) into the above equation and solving w.r.t. $\mathbb{E}(\mathbf{Q}_{k_1, i_1} \mathring{\mathbf{Q}}_{k_2, i_2} \mathring{\mathbf{Q}}_{k_3, i_3})$, we obtain that

$$\begin{aligned} \mathbb{E}(\mathbf{Q}_{k_1, i_1} \mathring{\mathbf{Q}}_{k_2, i_2} \mathring{\mathbf{Q}}_{k_3, i_3}) &= \mathbf{R}_{k_1, k_1} \delta(k_1 - i_1) \mathbb{E}(\mathring{\mathbf{Q}}_{k_2, i_2} \mathring{\mathbf{Q}}_{k_3, i_3}) \\ &\quad + \frac{\sigma^2}{t} \mathbb{E} \left[(\mathbf{QDQ})_{k_1, i_2} (\mathbf{QY}\tilde{\mathbf{D}}\tilde{\mathbf{R}}\mathbf{Y}^H \mathbf{R})_{k_2, i_1} \mathring{\mathbf{Q}}_{k_3, i_3} \right] \\ &\quad + \frac{\sigma^2}{t} \mathbb{E} \left[(\mathbf{QDQ})_{k_1, i_3} (\mathbf{QY}\tilde{\mathbf{D}}\tilde{\mathbf{R}}\mathbf{Y}^H \mathbf{R})_{k_3, i_1} \mathring{\mathbf{Q}}_{k_2, i_2} \right] \\ &\quad - \sigma^2 \mathbb{E} \left[\mathring{\beta} (\mathbf{QY}\tilde{\mathbf{D}}\tilde{\mathbf{R}}\mathbf{Y}^H \mathbf{R})_{k_i, i_1} \mathring{\mathbf{Q}}_{k_2, i_2} \mathring{\mathbf{Q}}_{k_3, i_3} \right] \end{aligned}$$

Writing $\mathbb{E}(\mathbf{Q}_{k_1, i_1} \mathring{\mathbf{Q}}_{k_2, i_2} \mathring{\mathbf{Q}}_{k_3, i_3})$ as

$$\mathbb{E}(\mathbf{Q}_{k_1, i_1} \mathring{\mathbf{Q}}_{k_2, i_2} \mathring{\mathbf{Q}}_{k_3, i_3}) = \mathbb{E}(\mathbf{Q}_{k_1, i_1}) \mathbb{E}(\mathring{\mathbf{Q}}_{k_2, i_2} \mathring{\mathbf{Q}}_{k_3, i_3}) + \mathbb{E}(\mathring{\mathbf{Q}}_{k_1, i_1} \mathring{\mathbf{Q}}_{k_2, i_2} \mathring{\mathbf{Q}}_{k_3, i_3})$$

and using (8.73), we obtain that

$$\begin{aligned} \mathbb{E}(\mathring{\mathbf{Q}}_{k_1, i_1} \mathring{\mathbf{Q}}_{k_2, i_2} \mathring{\mathbf{Q}}_{k_3, i_3}) &= \frac{\sigma^2}{t} \mathbb{E} \left[(\mathbf{QDQ})_{k_1, i_2} (\mathbf{QY}\tilde{\mathbf{D}}\tilde{\mathbf{R}}\mathbf{Y}^H \mathbf{R})_{k_2, i_1} \mathring{\mathbf{Q}}_{k_3, i_3} \right] \\ &\quad + \frac{\sigma^2}{t} \mathbb{E} \left[(\mathbf{QDQ})_{k_1, i_3} (\mathbf{QY}\tilde{\mathbf{D}}\tilde{\mathbf{R}}\mathbf{Y}^H \mathbf{R})_{k_3, i_1} \mathring{\mathbf{Q}}_{k_2, i_2} \right] \\ &\quad - \sigma^2 \mathbb{E} \left[\mathring{\beta} (\mathbf{QY}\tilde{\mathbf{D}}\tilde{\mathbf{R}}\mathbf{Y}^H \mathbf{R})_{k_i, i_1} \mathring{\mathbf{Q}}_{k_2, i_2} \mathring{\mathbf{Q}}_{k_3, i_3} \right] + o\left(\frac{1}{t^2}\right) \end{aligned} \tag{8.88}$$

We consider a unit norm deterministic row vector \mathbf{u} and define $\eta = \mathbf{uQ}\mathbf{u}$, $\rho_1 = \mathbf{uQDQ}\mathbf{u}^H$ and $\rho_2 = \mathbf{uQY}\tilde{\mathbf{D}}\tilde{\mathbf{R}}\mathbf{Y}^H \mathbf{R}\mathbf{u}^H$. Using (8.88), we get that

$$\mathbb{E}(\mathring{\eta}^3) = \frac{2\sigma^2}{t} \mathbb{E}(\rho_1 \rho_2 \mathring{\eta}) + \sigma^2 \mathbb{E}(\mathring{\beta} \rho_2 \mathring{\eta}^2)$$

We write $\mathbb{E}(\rho_1 \rho_2 \mathring{\eta})$ as

$$\mathbb{E}(\rho_1 \rho_2 \mathring{\eta}) = \mathbb{E}(\rho_1) \mathbb{E}(\mathring{\rho}_2 \mathring{\eta}) + \mathbb{E}(\rho_2) \mathbb{E}(\mathring{\rho}_1 \mathring{\eta}) + \mathbb{E}(\mathring{\rho}_1 \mathring{\rho}_2 \mathring{\eta})$$

$\mathbb{E}(\rho_1)$ is uniformly bounded while $\mathbb{E}(\mathring{\rho}_2 \mathring{\eta})$ is a $O(\frac{1}{t})$ term. $\mathbb{E}(\rho_2) \mathbb{E}(\mathring{\rho}_1 \mathring{\eta})$ is a $O(\frac{1}{t})$ term for the same reasons. Finally, we remark that $|\mathring{\rho}_1| \leq \frac{2d_{max}}{\sigma^4}$. Therefore, $\mathbb{E}(\mathring{\rho}_1 \mathring{\rho}_2 \mathring{\eta})$ is a $O(\frac{1}{t})$ term, and $\frac{2\sigma^2}{t} \mathbb{E}(\rho_1 \rho_2 \mathring{\eta})$ is a $O(\frac{1}{t^2})$ term.

In order to evaluate $\mathbb{E}(\mathring{\beta} \rho_2 \mathring{\eta}^2)$, we write

$$\mathbb{E}(\mathring{\beta} \rho_2 \mathring{\eta}^2) = \mathbb{E}(\rho_2) \mathbb{E}(\mathring{\beta} \mathring{\eta}^2) + \mathbb{E}(\mathring{\beta} \mathring{\rho}_2 \mathring{\eta}^2)$$

$\mathbb{E}(\rho_2)$ is uniformly bounded. $\mathbb{E}(\mathring{\beta} \mathring{\eta}^2) = O(\frac{1}{t^2})$ holds by the Schwartz inequality. We finally write that

$$|\mathbb{E}(\mathring{\beta} \mathring{\rho}_2 \mathring{\eta}^2)| \leq \left(\mathbb{E}|\mathring{\rho}_2|^2 \right)^{1/2} \left(\mathbb{E}|\mathring{\beta}|^4 \right)^{1/4} \left(\mathbb{E}|\mathring{\eta}|^8 \right)^{1/4}$$

and use Lemma 5 to justify that $\mathbb{E}(\beta_{\rho_2}^{\circ\circ}\eta^2) = O(\frac{1}{t^2})$. This completes the proof of (8.19).

8.8 Proof of Lemma 1.

We first establish that

$$\mathbb{E}(\log(1 + \epsilon_j))^2 < C \quad (8.89)$$

for some constant C independent of j and t . For this, we remark that

$$\sigma^2 \mathbf{Q}_{T,j,j} = \frac{1}{1 + \beta_j} \quad (8.90)$$

is less than 1. Therefore, $-\mathbb{E}(\log(\sigma^2 \mathbf{Q}_{T,j,j})) \geq 0$. As $\log(1 + \epsilon_j)$ is equal to $\log(\sigma^2 \mathbf{Q}_{T,j,j}) - \mathbb{E}(\log(\sigma^2 \mathbf{Q}_{T,j,j}))$, we get that

$$\log(1 + \epsilon_j) \geq \log(\sigma^2 \mathbf{Q}_{T,j,j}) = -\log(1 + \beta_j)$$

$\beta_j > 0$ implies that $\log(1 + \beta_j) \leq \beta_j$. Therefore, $\log(1 + \epsilon_j) \geq -\beta_j$ and $(\log(1 + \epsilon_j))^2 \leq (\beta_j)^2$. In order to prove (8.89), it is thus sufficient to establish that $\mathbb{E}(\beta_j^2) \leq C$. We denote by \mathbf{h}_j the column j of matrix \mathbf{H} . β_j is upperbounded by the match filter bound $\frac{\|\mathbf{h}_j\|^2}{\sigma^2}$. Using the expression of vector \mathbf{h}_j in terms of matrices $\mathbf{C}_R, \mathbf{C}_T$ and \mathbf{H}_{iid} , it is easy to check that

$$\mathbb{E}(\|\mathbf{h}_j\|^4) \leq C$$

for some constant C independent of j and t . This completes the proof of (8.89). Note that (8.89) implies that for each j , $\mathbb{E}|\log(1 + \epsilon_j)| < \infty$, a property which was implicitly assumed in the proof of Theorem 3.

We now complete the proof of Lemma 1. We consider a constant $A \in (0, 1)$, and express $\log(1 + \epsilon_j)$ as

$$\log(1 + \epsilon_j) = \sum_{k=1}^{\infty} (-1)^{k-1} \frac{\epsilon_j^k}{k} \mathbb{I}_{|\epsilon_j| < A} + \log(1 + \epsilon_j) \mathbb{I}_{|\epsilon_j| \geq A}$$

where for any set \mathcal{B} , $\mathbb{I}_{\mathcal{B}}$ is equal to 1 on \mathcal{B} and 0 outside \mathcal{B} . This leads to the following expression of $\mathbb{E}(r_j)$

$$\mathbb{E}(r_j) = \sum_{k=4}^{\infty} (-1)^{k-1} \mathbb{E} \left(\frac{\epsilon_j^k}{k} \mathbb{I}_{|\epsilon_j| < A} \right) - \mathbb{E}(\epsilon_j \mathbb{I}_{|\epsilon_j| \geq A}) + \frac{1}{2} \mathbb{E}(\epsilon_j^2 \mathbb{I}_{|\epsilon_j| \geq A}) - \frac{1}{3} \mathbb{E}(\epsilon_j^3 \mathbb{I}_{|\epsilon_j| \geq A}) + \mathbb{E}(\log(1 + \epsilon_j) \mathbb{I}_{|\epsilon_j| \geq A}) \quad (8.91)$$

Using (8.28) and (8.83), we remark that

$$\mathbb{E}|\epsilon_j|^8 \leq \frac{C}{t^4} \quad (8.92)$$

From the Markov inequality and the Holder inequality, we obtain that

$$P(|\epsilon_j| > A) \leq \frac{C}{t^4} \quad (8.93)$$

and

$$\begin{aligned} \mathbb{E}(|\epsilon_j|^6) &\leq \frac{C}{t^3} \\ \mathbb{E}(|\epsilon_j|^4) &\leq \frac{C}{t^2} \\ \mathbb{E}(|\epsilon_j|^3) &\leq \frac{C}{t^{3/2}} \\ \mathbb{E}(|\epsilon_j|^2) &\leq \frac{C}{t} \end{aligned} \quad (8.94)$$

By the Schwartz inequality,

$$|\mathbb{E}(\epsilon_j \mathbb{I}_{|\epsilon_j| \geq A})| \leq (P(|\epsilon_j| > A))^{1/2} (\mathbb{E}|\epsilon_j|^2)^{1/2}$$

(8.93) and (8.94) thus imply that $|\mathbb{E}(\epsilon_j \mathbb{I}_{|\epsilon_j| \geq A})|$ is upperbounded by $\frac{C}{t^{5/2}}$. We obtain similarly that

$$\mathbb{E}(\epsilon_j^2 \mathbb{I}_{|\epsilon_j| \geq A}) \leq \frac{C}{t^3}$$

and

$$\mathbb{E}(\epsilon_j^3 \mathbb{I}_{|\epsilon_j| \geq A}) \leq \frac{C}{t^{7/2}}$$

Using (8.89), (8.93) and the Schwartz inequality yields

$$|\mathbb{E}(\log(1 + \epsilon_j) \mathbb{I}_{|\epsilon_j| \geq A})| \leq \frac{C}{t^2}$$

We now establish that

$$\sum_{k=4}^{\infty} \mathbb{E} \left(\frac{|\epsilon_j|^k}{k} \mathbb{I}_{|\epsilon_j| < A} \right) \leq \frac{C}{t^2} \quad (8.95)$$

For $k \geq 4$, $\mathbb{E}(|\epsilon_j|^k \mathbb{I}_{|\epsilon_j| < A}) \leq A^{k-4} \mathbb{E}(|\epsilon_j|^4)$. Therefore,

$$\sum_{k=4}^{\infty} \mathbb{E} \left(\frac{|\epsilon_j|^k}{k} \mathbb{I}_{|\epsilon_j| < A} \right) \leq \left(\sum_{k=0}^{\infty} \frac{A^k}{k+4} \right) \mathbb{E}(|\epsilon_j|^4)$$

As $A < 1$, $\sum_{k=0}^{\infty} \frac{A^k}{k+4} < \infty$ so that (8.95) follows from (8.94). Putting all the pieces together gives $|\mathbb{E}(r_j)| \leq \frac{C}{t^2}$.

Chapitre 9

Conclusion

Les travaux présentés dans ce document sont dédiés à l'étude des performances en downlink des systèmes MIMO équipés d'un récepteur linéaire de type MMSE à l'aide de résultats issus de la théorie des matrices aléatoires. Pour différentes hypothèses sur les systèmes considérés nous déterminons des approximations asymptotiques en grande dimension du SINR en sortie du récepteur MMSE ainsi que de la capacité ergodique correspondante et en montrons la pertinence pour l'étude de systèmes réalistes. Nous utilisons alors ces approximations afin de simplifier l'étude et la résolution de problèmes d'allocations de puissance et de conceptions de matrices de précodage visant à maximiser la capacité ergodique.

Dans un premier temps nous nous intéressons au cas des systèmes MIMO multi-utilisateurs et multi-cellulaire ayant recours au CDMA comme technique d'accès multiple. En considérant un modèle des signaux proche de celui utilisé dans la norme WCDMA et un modèle de canal de propagation sélectif en fréquence présentant une matrice de covariance à l'émetteur différente de l'identité, nous généralisons les résultats précédemment obtenus dans l'article de conférence [7] dédiés à l'étude des performances des systèmes MIMO multi-utilisateurs mono-cellulaire et dans lequel nous avons montré que dans le cas d'une covariance à l'émission diagonale, une stratégie d'allocation de puissance uniforme n'est pas forcément optimale mais qu'une stratégie de sélection d'antennes est préférable. Ces résultats contrastent fortement avec ceux obtenus dans le cas des canaux i.i.d. pour lesquels le meilleur schéma de transmission maximisant la capacité de Shannon se trouve être une allocation de puissance uniforme parmi l'ensemble des antennes de transmission. Les résultats de simulation présentés permettent de s'assurer de la pertinence de l'étude en grande

dimension même pour des systèmes réalistes dotés d'un faible nombre d'utilisateurs et d'un faible nombre d'antennes. A ce sujet nous pouvons constater que la présence de stations de bases interférentes contribue à améliorer la vitesse de convergence des SINRS mis en jeu vers leur valeur moyenne. Afin d'illustrer l'utilité de ces développements théoriques, l'approximation de la capacité ergodique obtenue est finalement utilisée dans le but de déterminer une méthode d'allocation de puissance à long terme visant à garantir à l'ensemble des utilisateurs de la cellule un taux moyen d'erreur cible.

Ensuite, notre étude se focalise sur les systèmes de communication MIMO génériques dotés d'un modèle de canal respectant une structure de Kronecker. Dans [13], en considérant un canal sélectif en fréquence sujet à un éventuel phénomène de corrélation spatiale au niveau de l'émetteur, nous avons pu fournir une approximation en grande dimension de la capacité ergodique dont le terme d'erreur relatif associé se trouve être un terme d'ordre $O\left(\frac{1}{t}\right)$. Nous avons alors montré que celle-ci est valable pour un faible nombre d'antenne et donc potentiellement exploitable dans un contexte d'utilisation réelle. Nous avons déterminé la pertinence des précodeurs $\mathbf{K} = \mathbf{U}\mathbf{D}^{-\frac{1}{2}}\mathbf{\Lambda}_{opt}^{1/2}$ où \mathbf{U} et \mathbf{D} sont les matrices des vecteurs et valeurs propres associés à la matrice de covariance à l'émetteur et où $\mathbf{\Lambda}_{opt}$ est une matrice diagonale positive dont les entrées sont solution d'un certain problème d'optimisation faisant intervenir l'approximation en grande dimension de la capacité ergodique et pouvant être résolu à l'aide d'un simple algorithme du gradient. Nous y montrons finalement l'intérêt de cette approche par rapport à une optimisation directe de la capacité ergodique réelle dont l'expression compliquée entraîne la nécessité de recourir à des techniques de simulations de Monte Carlo extrêmement coûteuses, les deux méthodes fournissant des résultats très proches.

Finalement, constatant une légère faiblesse de l'approximation précédente lorsque le nombre de trajets associés au canal de propagation est égal à 1, ce qui correspond à un canal plat, nous avons entrepris de l'améliorer dans l'article de conférence [19] puis dans l'article de journal [14]. Nous avons ainsi considéré un canal de propagation plat doublement corrélé selon un modèle de Kronecker et dérivé une nouvelle approximation de la capacité ergodique dont le terme d'erreur associé est cette fois-ci un terme en $O\left(\frac{1}{t^2}\right)$. La vitesse de convergence de cette nouvelle approximation étant plus rapide, nous avons ainsi pu montrer sa pertinence pour l'étude des performances des systèmes MIMO dans un environnement de propagation non sélectif

en fréquence. Nous nous sommes une nouvelle fois assuré de la pertinence des pré-codeurs $\mathbf{K} = \mathbf{U}\mathbf{D}^{-\frac{1}{2}}\mathbf{\Lambda}_{opt}^{1/2}$ où les entrées de la matrice diagonales positives $\mathbf{\Lambda}_{opt}$ sont solutions d'un problème d'optimisation faisant intervenir cette nouvelle approximation de la capacité ergodique. Dans le cas d'une covariance diagonale à l'émission le problème précédent pouvant s'écrire analytiquement, nous avons pu montrer que la meilleur stratégie d'allocation de puissance par antenne n'est pas nécessairement une allocation de puissance uniforme mais qu'une stratégie de sélection d'antennes est plus indiqué.

Afin de compléter les résultats obtenus durant cette thèse, il serait certainement intéressant de déterminer une approximation en grande dimension de la capacité ergodique des systèmes MIMO dans le cadre de canaux de propagation sélectif en fréquence et doublement corrélé. Celle-ci pourrait alors être utilisée dans le but de déterminer les entrées de matrices de précodage permettant d'atteindre les débits limites admissibles.

Bibliographie

- [1] E. Telatar. Capacity of multi-antenna gaussian channels. *Europ. Trans. Telecom.*, 10 :585–595, November 1999.
- [2] R.R. Muller. A random matrix model of communication via antenna arrays. *IEEE Trans. on Information Theory*, 48 :2495–2506, September 2002.
- [3] J.M. Kahn C.-N. Chuah, D.N.C. Tse and R.A. Valenzuela. Capacity scaling in mimo wireless systems under correlated fading. *IEEE Trans. on Information Theory*, 48 :637–650, Mar 2002.
- [4] A. Lozano A.M. Tulino and S. Verdu. Impact of antenna correlation on the capacity of multiantenna channels. *IEEE Trans. on Information Theory*, 51 :2491–2509, July 2005.
- [5] G. Taricco E. Biglieri. Transmission and reception with multiple antennas : Theoretical foundations. *Foundations and Trends in Communications and Information Theory*, 2, 2004.
- [6] H. Viswanathan A. Mantravadi, V.V. Veeravalli. Spectral efficiency of mimo multiaccess systems with single-user decoding. *IEEE J. on Selected Areas on Com.*, 21, April 2003.
- [7] B. Mouhouche C. Artigue, P. Loubaton. On the performance of spatial-multiplexing mimo-cdma downlink receivers based on chip rate equalization and despreading. In *IEEE International Conference on Acoustics, Speech and Signal Processing, Las Vegas, USA*, March 2008.
- [8] S. Barbarossa G. Giannakis H. Sampath A. Scaglione, P. Stoica. Optimal design for space-time linear precoders and decoders. *IEEE Trans. on Signal Processing*, 50 :1051–1064, May 2002.
- [9] A. Paulraj H. Bolcskei, D. Gesbert. On the capacity of ofdm-based spatial multiplexing systems. *IEEE Trans. on Communications*, 50 :225–234, February 2005.

- [10] J-T. Chen C-K. Wen, P. Ting. Asymptotic analysis of mimo wireless systems with spatial correlation at the receiver. *IEEE Trans. on Communications*, 54 :349–363, February 2006.
- [11] S. Lasaulce J. Dumont, P. Loubaton. On the capacity achieving covariance matrix for rician mimo channels : an asymptotic approach. In *Proc. Globecom 2006*, December 2006.
- [12] E. Riegler G. Taricco. On the ergodic capacity of the asymptotic separately-correlated rician fading mimo channel with interference. in *Proc. ISIT 2007, Nice, France*, pages 531–535.
- [13] P. Loubaton C. Artigue. On the ergodic capacity of frequency selective mimo systems equipped with mmse receivers : An asymptotic approach. In *IEEE Global Communications Conference, New Orleans, USA*, November 2008.
- [14] P. Loubaton C. Artigue. On the precoder design of flat fading mimo systems equipped with mmse receivers : a large system approach. *IEEE Transactions on Information Theory*, June 2009.
- [15] A.M. Sengupta A.L. Moustakas, S.H. Simon. Mimo capacity through correlated channels in the presence of correlated interference and noise : A (not so) large n analysis. *Trans. on Inf. Theo.*, 49 :2545–2561, October 03.
- [16] P. Loubaton-J. Najim L. Pastur W. Hachem, O. Khorunzhiy. A new approach for capacity analysis of large dimensional multi-antenna channels. *IEEE Trans. on Information Theory*, 54 :3987–4004, September 2008.
- [17] A.L. Moustakas K.R. Kumar, G. Caire. Asymptotic performance of linear receivers in mimo fading channels. *IEEE Trans. on Information Theory*, 55 :4398–4418, October 2009.
- [18] G. Caire A.L. Moustakas, R. Kumar. Performance of mmse mimo receivers : a large n analysis for correlated channels. in *Proc. VTC 2009*, April 2009.
- [19] P. Loubaton C. Artigue. On the ergodic capacity and precoder design of flat fading mimo systems equipped with mmse receivers. In *IEEE International Symposium on Information Theory, Seoul, Korea*, June 2009.
- [20] C. Artigue B. Mouhouche. Coding strategies for cdma packet data networks with reduced rank multiuser detection, arq and packet combining. In *IEEE International Workshop on Signal Processing Advances in Wireless Communications, Recife, Brazil*, July 2008.

- [21] 3GPP. *TS 25.212, Multiplexing and channel coding (FDD)*.
- [22] 3GPP. *TS 25.213, Spreading and Modulation (FDD)*.
- [23] R.H. Clarke. A statistical theory of mobile radio reception. *Bell Sys. Tech. Journal*, 47 :987–1000, 1968.
- [24] J.M. Chauffray. *Detection et demodulation de stations de base dans un reseau UMTS*. PhD thesis, Universite de Marne-la-Vallee, 2003.
- [25] A. Toskala H. Holma, editor. *WCDMA for UMTS : Radio Access for Third Generation Mobile Communications*. Wiley Technology Publishing, 2000.
- [26] C. Yang S. *CDMA RF System Engineering*. Artech House, 1998.
- [27] E. Visotsky C.D. Franck. Adaptive interference suppression for cdma with long spreading codes. In *Allerton Conf., Monticello, Illinois, USA*, September 1998.
- [28] D.T.M. Slock M. Lenardi, A. Medles. A sinr maximizing rake receiver for ds-cdma downlinks. In *34th Asilomar Conf., Pacific Grove, California*, October 2000.
- [29] S. Verdu H. Poor. Probability of error in mmse multiuser detection. *IEEE Transactions on Information Theory*, 43 :858–871, May 1997.
- [30] D. Tse J. Zhang, E.K.P. Chong. Output mai distributions of linear mmse multiuser receivers in ds-cdma systems. *IEEE Transactions on Information Theory*, 47 :1128–1144, March 2001.
- [31] N. Jindal-S. Vishwanath A. Goldsmith, S.A. Jafar. Capacity limits of mimo channels. *IEEE J. Sel. Areas in Comm.*, 21, June 2003.
- [32] S. A. Jafar and A. Goldsmith. Multiple-antenna capacity in correlated rayleigh fading with channel covariance information. *IEEE Transactions on Wireless Communications*, 4 :990–997, May 2005.
- [33] A. Tulino A. Lozano and S. Verdú. Capacity-achieving input covariance for single-user multi-antenna channels. *IEEE Trans. Wireless Communications*, 5 :662–671, March 2006.
- [34] S.H. Simon A.L. Moustakas. On the outage capacity of correlated multipath mimo channels. *IEEE Trans. on Inf. Theo.*, 53 :3887–3903, November 2007.
- [35] A. Paulraj M. Vu. Capacity optimization for rician correlated mimo wireless channels. in *Proc. Asilomar Conference*, pages 133–138, November 2005.

- [36] S. Verdu A.M. Tulino. Random matrix theory and wireless communications. *Foundations and Trends in Communications and Information Theory*, 1, 2004.
- [37] J.M. Kahn-R.A. Valenzuela C.N. Chuah, D.N.C. Tse. Capacity scaling in mimo wireless systems under correlated fading. *IEEE Trans. Inf. Theo.*, 48 :637–650, March 2002.
- [38] P. Loubaton C. Artigue. On the ergodic capacity and precoder design of flat fading mimo systems equipped with mmse receivers. *In Proc. ISIT 2009*, July 2009.
- [39] W. Hachem A. Kammoun, M. Kharouf and J. Najim. Ber and outage probability approximations for lmmse detectors on correlated mimo channels. *IEEE Trans. on Information Theory*, 55 :4386–4397, October 2009.
- [40] S. Verdu, editor. *Multiuser Detection*. Cambridge University Press, New York : 1998.
- [41] L.A. Pastur. A simple approach for the study of global regime of large random matrices. *Ukrainian Mathematical Journal*, 57 :936–966, June 2005.
- [42] S. Lasaulce-P. Loubaton J. Najim J. Dumont, W. Hachem. On the capacity achieving covariance matrix for rician mimo channels : an asymptotic approach. *IEEE Transactions on Information Theory*.